X-SIGNAL AMPLIFICATION BY RUNT MEDIATED ANTAGONISM OF GROUCHO IN *DROSOPHILA*

A Dissertation

by

SHARVANI MAHADEVARAJU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	James W. Erickson
Committee Members,	Paul E. Hardin
	Gary R. Kunkel
	Keith A. Maggert
Head of Department,	Thomas D. McKnight

August 2015

Major Subject: Biology

ABSTRACT

Cell fate decisions are made in response to small quantitative differences in signal molecules. However the mechanisms by which small differences in protein levels regulate all-or-nothing responses remain mysterious. Sex determination in *Drosophila* is an example where small, two-fold differences in X-Signaling elements lead to distinct developmental fates. The goal my study is to understand the molecular mechanism that distinguishes the two fold difference in X-signal. The *Sex-lethal* (*Sxl*) gene stands on top of the sex determination pathway of *Drosophila melanogaster*. When active, *Sxl* triggers female development by regulating downstream target genes, and when inactive, it allows male development by default. The primary sex determination signal is the number of X chromosomes. Four X-linked genes *sisA*, *scute*, *upd* and *runt*, known generally as X signal elements (XSEs), signal X chromosome dose to *Sxl*. In XX embryos the two-X dose of XSEs activates the female specific establishment promoter, *SxlPe* and the protein produced initiates an auto regulatory splicing loop that maintains Sxl expression thereafter. The single dose of XSEs in XY embryos is insufficient to activate *Sxlpe*, thus no *Sxl* protein is made and male development follows by default. We previously proposed a 'XSE-signal' amplification mechanism that hinges on the maternally supplied co-repressor Groucho (Gro), a member of the Gro/TLE1 family. In the absence of Gro, *SxlPe* is expressed in both the sexes in direct proportion to X chromosome dose, meaning there is no signal amplification in the absence of Gro. Antagonism of Gro function in females is a key part of or proposed signal amplification

scheme, but the mechanism of damping Gro function were hypothetical. Here I propose and test the hypothesis that the XSE gene *runt* functions to dampen Gro dependent repression and is thus a key part of signal amplification. Antagonism between Runt and Gro would likely be mediated through Runt's WRPY motif, a tetrapeptide that mediates Runt and Gro interactions throughout the animal kingdom. Studies with a variety of transgenes with modifications in the WRPY motif revealed that, motif is required for Runt to activate *Sxl* supporting the hypothesis that Runt interferes with Gro repression at *SxlPe*. My findings also suggest that Runt is distinct from other XSEs as it is not required for initial activation of *Sxl* but required to maintain *Sxl* activity in regions of the embryo carrying active Gro. The *runt* protein normally binds as a heterodimer with either the *Big brother (Bgb)* or *Brother (Bro)* proteins. I created *Bro* and *Bgb* double mutants to test the hypothesis that neither protein was required for Runt to activate *Sxl*. Unexpectedly, I found that either Bro or Bgb is required for oogenesis thus I was unable to draw conclusions about a requirement for Bro or Bgb in sex determination. I showed, in contradiction to published studies, that the *Bro* protein is dispensable for viability and fertility. These findings provide significant knowledge to understand the mechanism by which small differences in protein levels regulate cell fate decisions.

DEDICATION

To all the Graduate students who had the perseverance to complete the task by facing all the challenges at school and life.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Jim Erickson for his guidance throughout my graduate studies. I would also like to thank my committee members Dr. Keith Maggert, Dr. Paul Hardin and Dr. Gary Kunkel for helping me in understanding and handling my project. I extend my thanks to members of my lab, family, friends and faculty in the department of Biology at Texas A&M University who directly and indirectly encouraged me to pursue my studies.

TABLE OF CONTENTS

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Fundamental cell fate decisions are made in response to signals. Often these signals act for a brief period, but they can cause long lasting or even permanent changes in cell or developmental fates. In many cases distinct cell fates are achieved in response to very small quantitative differences in the concentrations of signaling molecules. In *Drosophila*, anterior-posterior, and dorsal-ventral patterning is defined by small and dynamic quantitative differences in signals. During anterior-posterior patterning, a gradient of transcription factor Bicoid defines cell fates whereas for the dorsal-ventral axis the specific nuclear concentration of Dorsal protein determines regional identity [1] [2]. Typically differences in signal concentrations are small, yet the responses are precise suggesting that the signals are in some way amplified. The Bicoid signal is believed to be amplified by cooperative binding of Bicoid at certain concentrations and by interactions with other transcription factors [3]. Similarly, interpretation of nuclear Dorsal concentrations depends on both binding site affinity and interactions with several different regionally localized proteins [4].

Sex determination systems often depend on interpreting small differences in signals. In *C. elegans* sexual fate is defined by a twofold quantitative difference in X- and autosomal-signal elements but accurate interpretation of the signal requires amplification at both transcriptional and post transcription regulatory steps [5] . *Drosophila* sex determination also depends on an accurate response to two-fold

differences in signaling elements. In flies, a twofold difference in the amount of X-linked signaling elements (XSEs) leads to opposite sexual fates. My thesis explores a possible mechanism of signal amplification during *Drosophila* sex determination. Before I describe my experiments and results I will introduce the fundamental aspects of the regulation of *Drosophila* sex determination and discuss our working model for how the sex determination signal is amplified to produce a reliable and accurate outcome. I will discuss the known molecular components of the signaling mechanism with a special emphasis on Runt and CBF-β proteins and the evidence that links them to signal amplification.

The primary sex determination signal in *Drosophila* is the number of X chromosomes [6] [7]. The embryos with two X chromosomes develop as females and embryos with a single X develop into males. The target of X chromosome signal is *Sex-lethal* (*Sxl*), the sex determination switch gene that stands on top of the sex determination pathway [8]. *Sxl* protein is produced only in females in response to two doses of XSEs signal regulates female development and dosage compensation. In males a single dose of XSEs is not enough signal and in the absence of *Sxl* protein males develop by default. Once produced in females *Sxl* protein production is maintained throughout the life of the fly by a positive feedback loop controlling *Sxl* mRNA splicing.

Establishment and maintenance of SXL by a feedback loop

Sxl has two distinct promoters a female-specific establishment promoter (*Pe*) and a maintenance promoter (*Pm*). *SxlPm* is transcribed in both sexes [9]. By default *SxlPm*-derived mRNAs contain a premature stop codon within the third exon, referred as 'male specific' exon, that when translated generates a truncated nonfunctional SXL. *SxlPm* preRNA; however is processed into functional mRNA in the presence of *Sxl* protein. SXL is a RNA binding protein with highly conserved RNA recognition domains that are present in the members of RNA binding protein family. There are high affinities binding sequences poly (U) for the binding of SXL on the introns flanking the third exon. SXL binds to these sites causing skipping of the third exon, and splicing in the 'female mode' and producing functional SXL protein. The initial source of SXL that initiates the splicing cascade is derived from the *SxlPe* promoter.

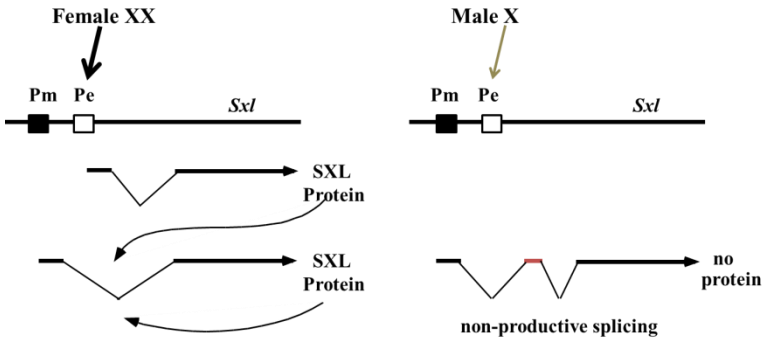

Fig.1.1. SXL auto regulatory feedback loop. *Sxl* has two promoters, the female specific, *SxlPe*, and *SxlPm* is active in both sexes. Activated in response to two X chromosomes, *SxlPe* produces an initial burst of *Sxl* protein. *Sxl* then maintain itself in the "on" state by autoregulatory processing of preRNA transcribed from *SxlPm* so that the translation terminating male specific exon is skipped. In males, the absence of initial burst of *SxlPe* causes preRNA from *SxlPm* to be spliced in the nonfunctional form.

Two doses of XSEs is enough signal to activate the establishment promoter (*Pe*) [10] [11] [12, 13] [14]. Transcripts from *SxlPe* are spliced directly from exon E1 to exon 4, skipping both exons 2 and 3 by default, so as to produce functional mRNAs. The early SXL protein produced by the transient activation of *SxlPe* initiates an auto regulatory feedback loop by splicing its own transcript [15] [16]. Once the autoregulatory loop is established there is no need for the production of SXL from *SxlPe* because SXL is thereafter maintained from the *SxlPm* promoter. Fig.1.1. The single dose of XSEs present on the X chromosome in males is insufficient to activate *SxlPe*. Thus even though *SxlPm* is transcribed, it fails to produce functional protein in the absence of initial burst of SXL from *SxlPe*.

SXL downstream targets and dosage compensation

Active production of SXL in females starts a gene regulatory cascade. SXL regulate its most immediate targets, *transformer* (*tra)*, and, *male specific lethal-2* (*msl*2), post transcriptionally. Fig.1.2 SXL directly regulates *tra* pre-mRNA splicing to produce functional TRA protein in females. In the absence of SXL in males default splicing leads to the addition of a stop codon that produces truncated TRA protein [17] [14]. TRA protein itself an RNA binding protein, regulates its downstream genes *dsx* and *fru* at the level of splicing. Both *dsx* and *fru* transcripts are alternatively spliced in the presence of TRA producing female specific isoforms, DSX^F and FRU^F. DSX^F regulates female development and differentiation by activating feminizing genes and repressing genes involved in male development and differentiation [18]. In males, where no TRA protein

is present both *dsx* and *fru* transcripts are spliced by default to produce male-specific DSX^M and FRU^M. DSX^M controls male development and differentiation by activating genes involved in male differentiation and also repressing genes that control female characters. FRU^M is known to control male courtship behavior [19] [20].

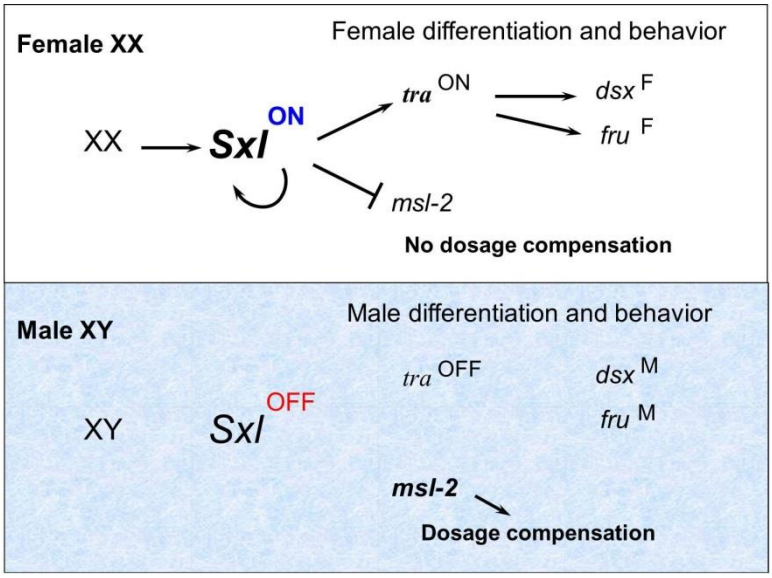

Fig. 1.2. **Sex determination hierarchy.** (Top panel) Embryos with two X chromosomes produce functional Sxl that maintains itself by an auto regulatory feedback loop. Sxl blocks *msl2* translation and directs production of functional *tra* protein. *tra* protein regulates production of female specific *dsx* and *fru*. (Bottom panel) *Sxl* is not produced in embryos with a single X. In the absence of *Sxl* and *tra* male specific *dsx* and *fru* mRNAs are produced by default, and msl2 mRNA is translated leading to dosage compensation.

The other known primary target of SXL is *msl2*. SXL inhibits *msl2* mRNA translation thus blocking production of a key component of the ribonucleoprotein complex (DCC) that controls dosage compensation. Dosage compensation in flies is

achieved by the elevation of expression from the single X chromosome in males to equal the level of expression from the two X in females [21] [22]. Thus in the presence of SXL the DCC does not form and the two X chromosomes are expressed at the typical female level. In the absence of SXL, MSL2 transcript is translated, and MSL2 assembles a functional DCC that elevates global expression from the male X chromosome.

X-signaling elements (XSEs)

X chromosome dose is specified at the molecular lever by at four zygotically expressed X-linked genes *scute* (*sisterless B*), *sisterless A* (*sisA*), *unpaired* (*upd* or *sisterless C also outstretched*), known generally as X-signal elements or XSEs [23] [24] [25] [13, 26] [27]. dMyc a bHLH factor encoded by *diminutive* (*dm*) has also been reported to be an X linked activator of *SxlPe* ; however, the *dm* gene is not expressed zygotically during the period when *SxlPe* is active , a result consistent with our lab's failure to duplicate the claimed synergistic genetic interactions between *dm* and any other strong XSEs [28, 29](unpublished data).

Scute and SisA are the strongest activators among XSEs required to initiate *SxlPe* expression in females. Both are among the earliest expressed genes in the fly befitting the early timing of *Sxl* activation. *Scute* encodes a basic helix loop helix (bHLH) transcription factor. *Scute* begin to transcribe in cycle 9 peaks in cycle 12-13 and ends abruptly at cyc 14 [24] [7]. Scute heterodimerizes with the maternally supplied bHLH protein, Daughterless (Da). The heterodimer is known to act typically as a direct activator by binding to *SxlPe* promoter [30] [31]. Being a prototypical heterodimer it

binds to canonical E-boxes in the promoter region but non-canonical binding sites are also important for biological function [30]. Genetic studies have confirmed that *da* is equally important as *scute* for the activation of *Sxl* as the loss of maternal *da* has the same drastic consequences as loss of zygotic *scute*.

SisA is a b-ZIP protein that is suspected to behave like a typical direct activator at *SxlPe*. *sis*A expression begins in nuclear division cycle 8. It is expressed uniformly throughout the cortex of the embryo peaking in cycles 12-13 and rapidly disappears from the cortex at cyc 14 [32]. White the *sisA* gene is silenced in somatic nuclei at cycle 14, it continues to be expressed in the internal yolk nuclei until much later [27]. Since SisA appears unable to from homodimers, it is thought to heterodimerize with a different protein. Our lab is still on a quest to find this partner.

Upd is a ligand for the JAK-STAT signaling pathway in contrast to the other XSEs which are transcription factors [33] [26] [25]. It begins to express in the syncytial blastoderm at nuclear division cycle 13. This is somewhat later than the expression of all the other XSEs and after the initiation of *SxlPe* itself. The effect of *upd* mutations effect on *Sxl* is weaker than all the other XSEs. Upd signals through Hopscotch (Hop, the Jak kinase) to activate the maternally supplied transcription factor Stat92E. There are three Stat binding sites on the *SxlPe* promoter. Loss of function *hop* alleles eliminates the *SxlPe* expression in the central region of the female embryo. Even though mutants of members JAK-STAT signaling pathway showed severe phenotype at cyc14 *SxlPe* came on normally at cyc 12 and 13. It suggests that *upd* and the JAK-STAT is involved in maintenance of *SxPe* at nuclear division cycle 13 and 14. It has been proposed that

having XSE elements target several steps in *SxlPe* activation is a means of increasing the fidelity and robustness of the sex determination switch [33].

Runt is a member of 'RUNX' – runt-related protein. *runt* is expressed at low levels as early as nuclear cycle 10 in a broad region of the embryo that excludes the anterior pole. At about nuclear cycle13, without forming any sharp borders *runt* expression is diminished towards both the poles. Thus Runt is unlike the other XSE genes in that its expression, and function, is spatially restricted. Slightly later, runt expression evolves into a typical 7 stripe pair rule pattern, but it is the broad low-level early expression pattern that regulates sex determination [34] [35]. Runx proteins function as dimers. The Runx subunit is sometimes known as CBF-α (Core binding factor) and its non-DNA binding partner is known as CBF-β [36]. The fly contains two CBF-β proteins: brother (Bgb) and Brother (Bro). Bgb is maternally contributed [37]. *Bro* is expressed for a brief period and expression of *Bro* closely mimics those of *scute* and *sisA* [37] [38]. It has been assumed that one or both CBF-β proteins work with Runt at *SxlPe*, but this has not yet been tested [39].

As mentioned earlier, XSEs work in conjunction with a variety of other proteins.Fig.1.3. Some like Da and STAT have obvious direct connects to XSE function whereas other proteins involved in sex determination may either facilitate or inhibit *Sxl* activation. A key inhibitor is the repressor *Deadpan* (*dpn*) a bHLH factor that binds directly to *SxlPe* [40] [41].[42] [43] [44]. Dpn expression begins prior to nuclear division cycle 12 and peaks in cycle 13. Dpn mRNA is accumulates to higher levels in the middle of the embryo than in the poles but is present everywhere [41]. Dpn is a Hes

family repressor containing a C-terminal 'WRPW' tetrapeptide motif known to interact with Groucho a Gro/TLE co-repressor [45] . Mutations in *dpn* cause ectopic *Sxl* expression in males, but the effect is considerably weaker than observed with loss of function mutations in the co-repressor Gro. This suggests that other bHLH repressors similar to Dpn might regulate *Sxl* through the same site. Hey is a HLH family protein that has an YRPW motif. It has relatively weak effect on *Sxl* expression and thought to repress *Sxl* by binding to Dpn-binding sites before the dpn protein is expressed [44]. Emc is a maternally supplied HLH repressor that lacks the basic DNA binding domain. It can dimerize with both Sc and Da but appears to play only a small role in regulating *SxlPe*.

The maternally contributed Groucho is member of the Gro/TLE1 member of co-repressors [45]. We have proposed that Gro is a key determinant in X-signal amplification because *Sxl* is expressed in direct relation to X-chromosome dose in *gro* mutants [44]. The mechanisms by which Gro repress *Sxl*, or its other targets are not known in detail. However, the consensus view is that Gro facilitates the formation of repressive chromatin at the target loci. It has been proposed to oligomerize and spread over long ranges to mediate repression over several kilobases but repression over short ranges, such as occurs at *SxlPe,* appears to be a more common phenomenon [46] [47]. Although we do not yet know how Gro silences genes, we do know that phosphorylation of Gro compromises its ability to repress its targets. In the early embryo, MAP kinase directly phosphorylates Gro at the poles of the embryo in response to Torso signaling [48]. The *torso*-controlled phosphorylation of Gro likely explains an early observation

that *Sxlpe* is uniformly expressed throughout the embryo in *runt* mutants if they also carry a *torso* gain of function mutation [23]. The most sensible explanation is that *runt* function is needed in the central regions of the embryo where Gro is fully active but is not required at the poles because Gro activity is repressed there. This suggests there is a different way of regulating *Sxl* at poles and there are regional differences in *Sxl* regulation with in the embryo.

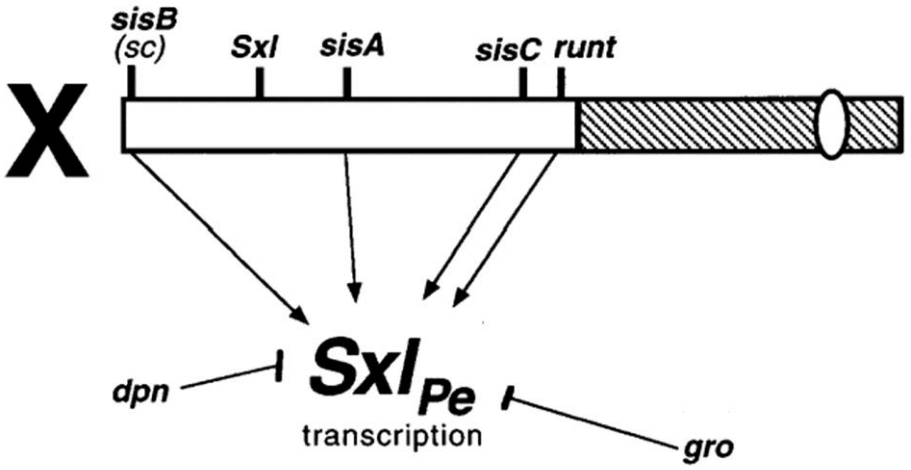

Fig.1.3. Factors that regulate *SxlPe*. X-linked signaling elements *sisA, sc, sisC* and *runt*. Autosomal repressor Dpn recruits maternal co-repressor Gro to *SxlPe*.

Nature of signal sensing at *SxlPe* and its response to several transcription factors

SxlPe is regulated by the cumulative action of several transcriptional activators and repressors. Critically it produces a transcript that is spliced by default to produce the functional protein that initiates the *Sxl* regulatory feedback loop. Transcription at *SxlPe* begins in nuclear division cycle 12 in most nuclei. It is highly active in all the nuclei

during cycle 13 and early in cycle 14. *SxlPe* activity ceases by 20 min into cycle 14 by which time *SxlPm* is active and autoregulatory splicing is established [7] [9].Fig.1.4.

SxlPe contain regulatory elements that respond to several transcription factors [43] [30]. Its regulatory region extends up to -3kb upstream of transcription start site of the early exon (E1). Transgenes containing 1.4kb of upstream sequence appear to drive almost wild type expression of *SxlPe-lacZ* fusions, and this region is sensitive to the dose of signaling elements and contains known cis regulatory elements for several of the key regulators. Transgenes carrying 0.4kb of upstream DNA express sex-specifically but expression is low level and patchy. There are 11 Scute/Da binding sites within -1.4kb region and six of them are clustered within -0.4kb [30]. SisA is thought to act through the -0.4kb region as 0.4kb *SxlPe-lacZ* fusions show sensitivity to *sisA* dose but in the absence of SisA dimerization partner no binding sites have been identified. There are three STAT sites within -1.4kb region -253, -393, and -428bp [33]. Even though *runt* is an activator of *SxlPe* no consensus binding sites exists on the promoter. Two canonical binding sites for the repressor Dpn, which recruits Gro to the promoter, are centered at -108, -119 and a non-canonical site at -160 [44].Fig.1.5.

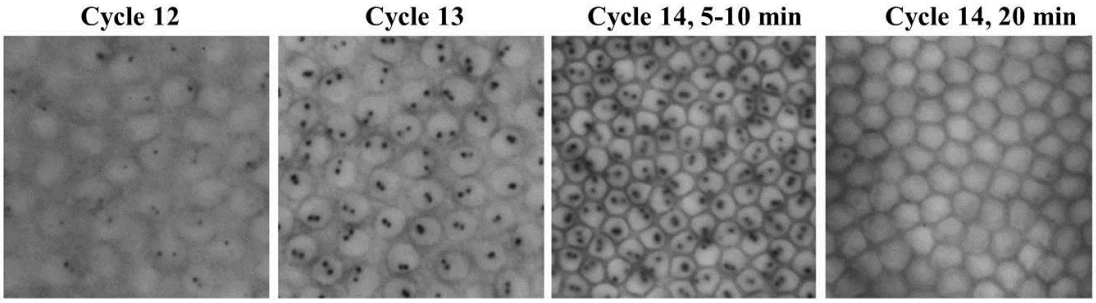

Fig.1.4. *In situ* hybridization to *Sxlpe* transcripts in WT embryo. Nascent transcripts are visible as dots in the surface view of embryo. Two dots represent transcription from X-linked *Sxlpe* in females that begin in nuclear cycle 12 continue until early cycle 14 and abort within 20 min into cycle 14.

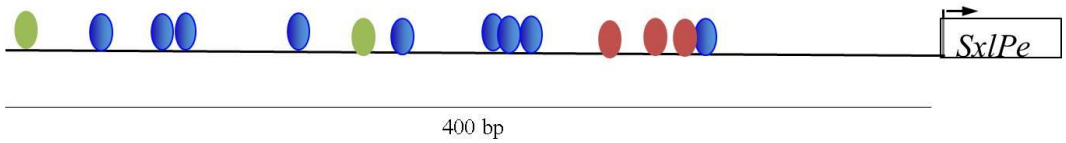

Fig.1.5. -0.4kb *Sxlpe* promoter representing known regulator binding sites. There are 9 Scute/Da sites (blue), 2 STAT sites (green) and 3 Dpn sites (green). SisA and Runt are proven to act through the promoter but no predicted binding sites are found.

An orchestra of transcriptional activators and repressors decides the sexual fate in fly. Those factors that are maternally provided or encoded on autosomes are present at the same concentrations in males and females. Only the XSE proteins differ in concentration, so ultimately the fly must distinguish between a two-fold difference in XSE protein concentrations to determine sexual fate decision. Our lab's working model for how this occurs posits that dose-sensitivity is achieved via a signal amplification

mechanism that requires inputs that increase activator function and that dampen repression.

X-dose signal amplification – a positive feedback mechanism

Previous experimental evidences suggested that signal amplification primarily depends on the corepressor, Groucho, a Gro/TLE1 member. In absence of Gro, *SxlPe* is ectopically expressed in males and is transcribed earlier in females. In the absence of Gro *SxlPe* is expressed in both the sexes in direct proportion to X chromosome dose, meaning that without Gro there is no X-signal amplification as illustrated by the model proposed by Lu, H., et al., in Fig.1.6 [44]. The essential features of the model are that Gro sets the initial threshold for *SxlPe* activation. Two X chromosomes provides sufficient XSE proteins to cross this activation threshold at nuclear cycle 12. Once initiated, the active state is maintained in by rising XSE levels from continued transcription of the XSE genes and by dampening Gro repression. In males, the single dose of XSEs does not provide enough XSE protein to cross the activation threshold set by the repressors so *SxlPe* is left inactive. The actions of Dpn and Gro continually adjust the repression threshold higher to compensate for the increasing XSE protein concentration in males as the nuclei double and redouble in cycles 13 and 14. The inhibition of Gro was proposed to work from the simple act of transcription increasing histone acetylation, which should itself antagonize Gro function; there is, however, no

direct evidence that this occurs. I have investigated an alternative means of inhibiting Gro function via the XSE protein Runt.

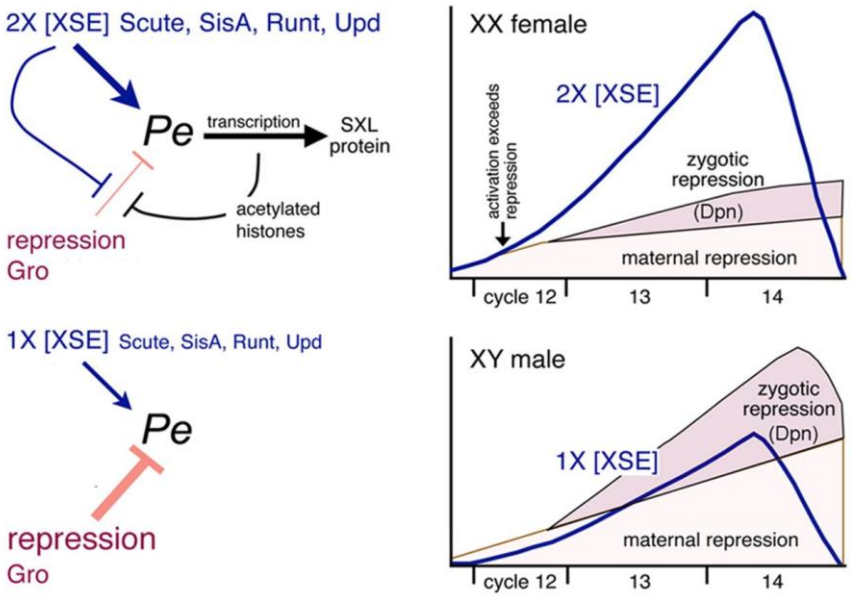

Fig.1.6. X-dose signal amplification.

Runt is the founding member of the 'RUNX' – Runt-related protein family transcription factor. Runx proteins heterodimerizes with a non DNA binding β subunit(CBF- β) [37]. There are three RUNX proteins in humans and all are transcription factors that regulate several developmental aspects [49]. Misexpression of Runx proteins has been shown to cause cancers and several other diseases. RUNX1 is most frequently involved in human acute leukemia, RUNX2 in Cleidocranial dysplasia

and RUNX3 in gastric cancer. Usually a chromosome rearrangement is responsible to cause the acute melanoid leukemia. Simple mutations in Runx gene or changes in the gene dose are also responsible for causing disease [50].

An interesting aspect of RUNX proteins is that they regulate their downstream targets context-dependently [51]. The decision of Runx to act as either an activator or repressor depends on the constitution of target gene regulatory regions at a specific time in a specific cell. Molecular mechanisms attribute to a specific sequence binding and interaction with different factors available in the cell at a specific time [52] [50] [53]. Most Runx protein function in mammals depends on their heterodimerization with CBF-β which is encoded by a single gene in mammals.

In *Drosophila* there are four 'RUNX' proteins *runt, lozenge, RunxA,* and *RunxB* [36]. The targets they regulate have been identified mainly by genetic approaches and the molecular mechanisms for most target genes are not well understood. Runt was first identified for its roles in segmentation in *Drosophila* [54]. Only later was it is learned that *runt* has functions as an XSE in sex determination and nervous system development [23].

Runt like all other members of Runx proteins is defined by a conserved 128 amino acid DNA binding 'runt domain' [55]. Runt binds to the core consensus site 'YGYGGY' or RCCRCR where Y and R represent pyrimidines and purines [36]. Runt proteins are also characterized by a C-terminal 'WRPY' tetrapeptide motif [56]. This motif is analogous to Hairy-enhancer-of-split (Hes) group of dedicated repressors. Hes proteins interact with co-repressor Gro through their C-terminal conserved 'WRPW' and

recruit Gro to target promoter to repress transcription. Similarly, the WRPY motif of Runx proteins interacts with Gro and is critical for repression by Runt in several developmental context [56]. WRPY motif is non-essential for either the DNA-binding properties or for their ability to dimerize with CBF- β proteins [56].

runt is a primary pair rule gene that regulates several segmentation genes [39]. It repress *even-skipped* (*eve*) at stripe 2, *engrailed(en), orthodenticle* (*otd*) posterior domain, *hairy* (*h*) stripes, and *Engrailed* (*En*). It activates *fushi tarazu* (*ftz*) and the sex determination switch gene *Sxl* at different times in development. One of the interesting regulation is that Runt either activates or represses *sloppy paired 1 (slp1)* and *wingless* depending on the cell types [57] [58].

Most studies of the molecular mechanisms of Runt function have relied on ectopic expression of *runt* or *runt* derivatives. Runt repression is mostly attributed to Gro interaction through the WRPY motif, but there are exceptions where Runt can repress independent of WRPY motif [56]. Runt's activation function is not well understood. Activation of one gene, *ftz*, is partially dependent on the WRPY motif [57]. Examples exist where activation and repression are independent of the Runt DNA binding domain and most regulatory decisions depend on other factors available in a specific cell at a specific time [50]. Table.1.1 is a summary of Runt's context dependent target genes and includes a note of what is known about mechanism.

Table 1.1. Summary of Runt target gene regulation.

Target Gene	Mode of Action	Known mechanism
ftz	Activation	DNA binding independent. WRPY independent (partially). Needs an orphan nuclear receptor FTZ-F1 that has recognition sites within fDE1 a cis-acting element within the 'Zebra' element of *ftz* promoter[56].
Sxl	Activation	DNA binding dependent [35].
eve	Repression of stripe 2	WRPY dependent [56].
hairy	Repression of stripe 1	WRPY dependent [56].
otd	Repression	WRPY independent. Repression also requires an orphan nuclear receptor, *tailless* [39].
wingless	Repression Activation	The combination of Runt and Ftz can acts either as activation complex or a repression complex depending on cell type. The regulatory regions might have specific binding sites for 'some other factor' that stabilize the combination into either one of them [57].
slp1	Repression Activation	WRPY independent. In anterior half of the even numbered para segments Runt requires Ftz for *slp1* repression. Ftz not only blocks Runt-dependent activation in these cells but the combination sufficient for *slp1* repression in all blastoderm nuclei [57]. Runt is required for *slp1* activation in odd-numbered parasegments. This Runt-dependent activation involves cooperation with the zinc-finger transcription factor encoded by the pair-rule gene *opa*. The simple combination of Runt + Opa is sufficient for *slp1*activation in all somatic blastoderm cells that 'do not' have Ftz [57].
Engrailed	Repression	There are two steps in repression. Establishment is independent of the Runt DNA binding domain and WRPY independent. Zn-finger proteins coded by *tramtrack* binds to *en* promoter and interact with Runt. Later maintenance step requires Runt's DNA binding domain and is WRPY dependent.. Hairless is suggested to be a factor in a common corepressor complex involving Gro and CtBP [58].

Drosophila contains two CBF-β proteins: Brother (Bro) and Big brother [38]. Bro and Bgb interact with Runt and confer high affinity DNA binding of the Runt subunit. Having conserved elements they interact in a cross-species manner. CBF-β is essential for mammalian Runt function and are shown to known to be required for several in vivo functions of Runt in flies [38] [37]. *Bgb* has been shown to enhance the mutant phenotypes of the Runx protein *Lozenge* in a sensitized mutant background during eye development proving its functional relation with Lozenge [59]. However *Bgb* null mutants alone do not have any defects in eye development leading to the idea that Bgb and Bro act redundantly. In addition when *Bro* was knocked down by RNA interference it produced a strong segmentation phenotype [60]. When *Bgb* alone was knocked down there was no effect on segmentation pattern; however, when Bgb and Bro were both inhibited an extreme segmentation phenotype was observed, further supporting the idea of redundancy between them [60]. There is evidence for specific separate function of Bgb in nervous system related to its unique expression pattern. Bgb and Bro have been assumed to work redundantly at *Sxl*, as both CBF-βs are present in the early embryo but there is no direct evidence for this. It is also possible that Runt functions independent of both the CBF-β in sex determination, a possibility suggested by the failure to find evidence for sequence-specific Runt binding at *SxlPe* [35]. (D. Yang and J. Erickson unpublished, see Chapter III). A CBF-β independent Runt function is not unprecedented of as there appears to be an example in mammals where Runx can function independent of CBF-β [61]. My experiments to address the question of whether Runt requires a CBF-β to regulate *Sxl* are addressed in Chapter III.

CHAPTER II

ROLE OF RUNT IN *SXL* REGULATION

runt is the founding member of the 'RUNX' or runt-related transcription factor family. It was originally identified for its pair-rule segmentation function in *Drosophila* [54]. Later, *runt* was identified as an X-signal element (XSE) for sex determination by virtue of its female-specific lethal genetic interactions with the XSE genes *sisA* and *sisB* and by the strong effects of *runt* null mutations on *Sxl* expression [23] [62].

Genetically *runt* is considered a weak XSE because alterations in *runt* gene dose cause much weaker lethal interactions with *sisA* and *scute* than are observed between the strong XSEs *sisA* and *scute*. In females, loss of one copy each of *sisA* and *scute* (example. sc^{-} + / + $sisA^{-}$) results in the death of nearly all female progeny, however, the synergism between *runt* and *scute* or *runt* and *sisA* is typically much weaker [23] [62]. In males, there is a strong synergistic lethal interaction between the two strong XSEs such that nearly all XY embryos bearing two copies of $sisA^{+}$ and $scute^{+}$ activate *Sxl* and do not survive, whereas, little or no lethality is seen with duplications of $runt^{+}$ and $sisA^{+}$ or $runt^{+}$ and $scute^{+}$ [23] [62].

While *runt* appears to be much less dose-sensitive than *sisA* or *scute*, it is nonetheless essential for proper activation of *Sxl* in XX embryos. Complete loss of *runt* function abolishes *SxlPe* activity and *Sxl* expression in a broad central region of female embryos but does not affect *Sxl* expression at the anterior or posterior poles [23]. Thus,

runt is unique among XSEs in that its functions are region-specific, and in that alterations in *runt* gene dose have much stronger effects in female than in male embryos.

In their initial characterization of *runt* as an XSE, Duffy and Gergen (1991) reported that gain-of-function mutations in the terminal *torso* signaling pathway that defines the anterior and posterior embryonic poles, suppressed the effects of *runt* mutations on *Sxl*. Specifically, mutations that caused *torso* kinase to be active in the entire embryo, rather than being restricted to the posterior and anterior poles, restored full uniform *Sxl* expression in *runt* null mutants [23]. At the time, this result was not easily explainable; however, later experiments have provided clues that the regulation of *Sxl* by *runt* and the control of the terminal system both likely involve the ubiquitous corepressor Groucho (Gro).

Groucho is a maternally contributed Gro/TLE1 member of co-repressors. Gro is recruited to *SxlPe* via repressor proteins such as Dpn [56] [44]. Where it defines the threshold XSE concentrations needed to activate *SxlPe* and maintains the promoter in the off state in males. In *gro* loss of function mutants, *SxlPe* is derepressed in males and comes on prematurely in females [44] [42] [41]. In *gro* mutants, *SxlPe* is expressed in both the sexes in direct proportion to X chromosome dose indicating that the XSE signal is not amplified in the absence of Gro [44]. Given these findings, our lab proposed that Gro repression is dampened in females as part of the XSE signal amplification mechanism, and in this Chapter I explore the role that Runt plays in depressing Gro function [44]. The idea that Runt might function to counteract Gro at *SxlPe* was originally proposed because Runt is known to bind Gro directly through Runt's

C-terminal tetrapeptide motif, WRPY but stronger correlative evidence has come from more recent studies showing that Gro's repression function is modulated at the embryonic poles via the Torso RTK signaling pathway [63] [48].

Gro is phosphorylated at the poles of the early embryo by MAP kinase in response to Torso signaling [48]. Since phosphorylated Groucho loses much of its repressive function, the early zygotic genes of the terminal system are expressed at the embryonic poles but repressed in the center of the embryo. Complete loss of *runt* function abolishes *Sxl* expression in a broad central region of female embryos but does not affect expression at the poles. This is strongly suggestive that Runt may be needed do dampen Gro repression in the central region of the embryo but be dispensable at the poles where Gro function is depleted because of the *torso* RTK signal.

The paradoxical, but appealing, notion is that Runt is acting as an activator of *SxlPe* through its co-repressor interaction motif. In most cases, corepressor binding would lead to repression of Runt targets, however, some aspect of the regulatory scheme at *SxlPe* likely causes Runt binding to inhibit Gro function. There may be precedent for such an interaction in human spleenogenesis where activation of two target genes by the homeodomain transcription factor, TLX1, appears to depend on an interaction with TLE1, a Gro homolog [64]. The research described in this chapter was designed to discover the mechanism by which Runt regulates *SxlPe*. My findings show that Runt is not needed for the initial activation of *Sxl*, but is instead required to maintain full expression in females. Activation requires Runt's WRPY motif or the related WRPW motif that is present in numerous dedicated repressor proteins. My findings are

consistent with the hypothesis that Runt activates *Sxl* primarily by local antagonism of Gro at *SxlPe*.

Runt is needed to maintain *SxlPe* activity, not to initiate it

I precisely monitored *SxlPe* transcription during early development by *in situ* hybridization in *runt*^{3} null mutant embryos. An intron-exon derived RNA probe detects *SxlPe* nascent transcripts in addition to mature transcripts. Nascent transcripts appear as dots in the nucleus and the two dots visible in most cycle 12 nuclei represent transcription initiation from both the X-linked *Sxl* alleles. During cycle 12 *runt*^{3} embryos were indistinguishable from WT.Fig.2.1A. Defects were first visible in cycle 13. In contrast to wild type, where *SxlPe* is active in all nuclei, *Sxl* expression begins to decline during cycle 13 in *runt*^{3} mutants. As expected, the decline is evident only in the central broad region but not at the poles. Fig.2.1A. By early cycle 14 (10-15min) expressions is completely extinguished in the broad central region in *runt*^{3} mutants compared to wild type that exhibits peak RNA levels during this period.Fig.2.1A and Fig.2.1B.

My observations suggest that *runt* is not required for the initial activation of *SxlPe*. Instead, *runt* is needed to keep the promoter active during cycles 13 and early 14. The key factor inhibiting *Sxl* expression during this period is Gro [44]. A plausible scenario is that Runt keeps *SxPe* active by antagonizing Gro repression. To test the

prediction that Runt interferes with Gro repression I modified Runt's WRPY motif in a series of transgenes and observed their effect on *Sxl*.

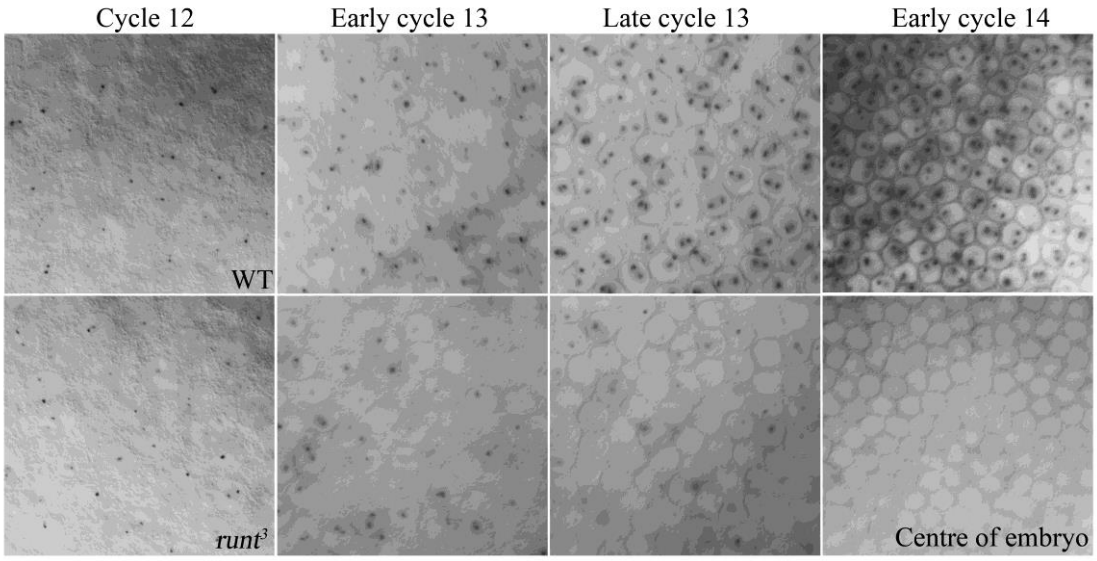

Fig.2.1A. *In situ* hybridization to nascent *SxlPe* transcripts. Surface views of embryos. Two dots represent transcription from the X-linked *Sxl* alleles in females. (Top panel) In wild type (WT) embryos transcription begins in cycle 12 in most nuclei. At the onset of cycle 13 all nuclei express *SxlPe* and continue to do so until early cycle 14. (Bottom panel) In *runt*^{3} embryos transcription begins in cycle 12 and is indistinguishable from WT. Defects are first apparent at the onset of cycle 13 and activity declines through cycle 13 except at the poles where expression equals wild type. *SxlPe* expression is completely extinguished from the broad central region in cycle 14.

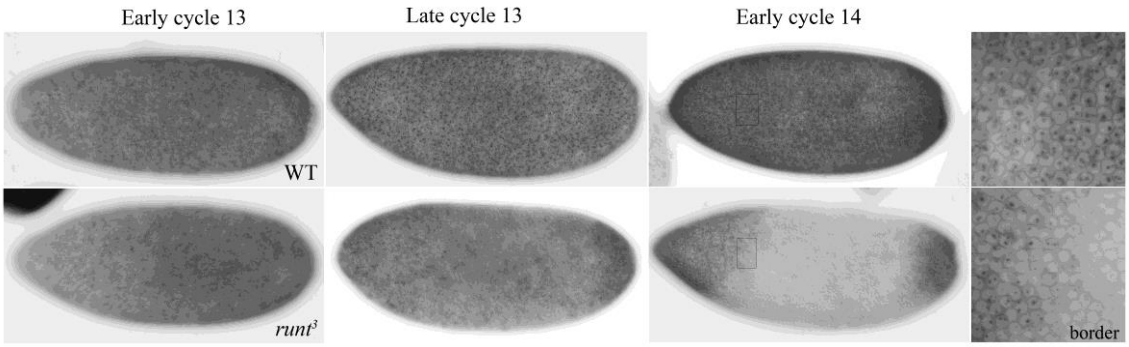

Fig.2.1B. *SxlPe* transcripts in whole embryos. Embryos are the same ones shown in Fig. 2.1A .(Top panel) Amount of *SxlPe* transcript increases in WT with time. (Bottom Panel) In *runt*^{3} mutants, *SxlPe* transcripts in the centers of the embryos gradually decline through cycle 13 and are eliminated by early cycle 14. Right hand panels show magnified surface views at the location of the anterior border between expressing and non-expressing cells in *runt*^{3} mutants. Embryos are oriented anterior to the left, dorsal to the top.

Generating *runt* transgene that provide XSE function

The *runt* gene has a large and complicated regulatory region so most of what is known of *runt's* function in flies is understood through over expression studies [65] [34]. Because the timing and levels of expression were critical to our analyses I wanted to create a transgene that provided full *runt* function for sex determination driven from its endogenous regulatory sequences. Using Klingler et al. as a guide I created a 10,050 bp transgene (from -4,702 bp upstream to +2,137 bp downstream of the *runt* transcription unit, details in methods) and integrated it using targeted φC31 mediated integration [34] [65].Fig.2.2A. I found that the wild-type transgene (named *runt·WRPY*) exhibited a normal early *runt* expression pattern studies as measured by in situ hybridization [34] [65].Fig.2.2B. *runt* expression was first detectable in nuclear cycle 10. Transcripts

gradually increased though cycles 13 without any visible expression in the anterior. By cycle 13 there was a broad high central region expression with greatly reduced mRNA in the posterior. The expression pattern from the transgene was indistinguishable from wild type and was confirmed by counting 100 embryos at cycle 13. While the early *runt* pattern, which is responsible for *runt's* sex determination function was normal, the later patterns were different from the endogenous *runt* pattern [35].

Normally *runt* expression occurs in 7 stripes during early cellularization and this resolves into a 14 stripe pattern at the beginning of gastrulation [65].Fig.2.2B. The early pair rule pattern was defective in the transgenes with only stripe 1 being visible at a time when all 7 stripes are seen in wild type. The transgenes do eventually produce all 7 stripes but expression level is comparatively low especially at the dorsal region of the embryo.Fig.2.2C.

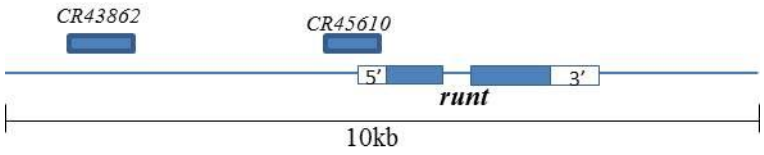

Fig.2.2A. *runt·WRPY* transgene. 10kb genomic DNA insert in *runt·WRPY* transgene. *CR4362* and *CR45610* are uncharacterized genes in the genome reside with the transgene.

Fig.2.2B. Early *runt* expression from transgenes (bottom row) is identical to the activation of the endogenous *runt* locus (WT top row). *runt* transcripts first appear in nuclear cycle 10. Transcripts accumulate throughout cycles 12 and 13. There is no detectable expression at the anterior poles and decreased expression in the posterior poles relative to the central embryo. Embryos are oriented anterior to the left, dorsal to the top.

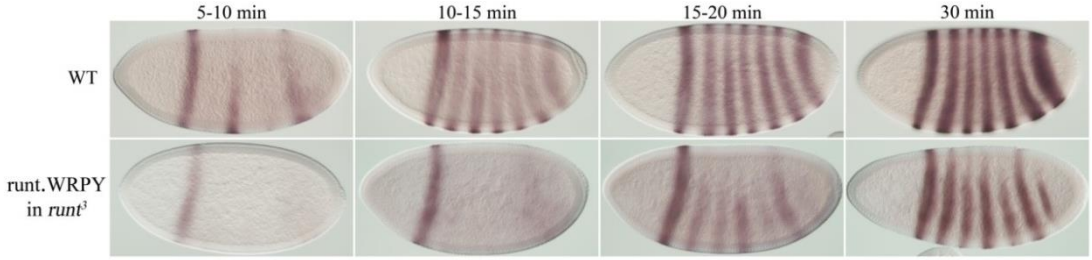

Fig.2.2C. *runt·WRPY* transgenes do not accurately mimic the pair rule expression pattern of the endogenous *runt* locus (WT) during cycle 14. Expression of *runt* changes dynamically from a broad to a striped pattern within 5-10 minutes of the onset of cycle 14. A single stripe develops in transgenes instead of three as in WT in 5-10 minutes, single stripe instead of seven in 10-15 minutes but all seven stripes appear between 15-30 minutes. Stripes are located as in wild-type, but are more weakly expressed, particularly in dorsal regions. Embryos are oriented anterior to the left, dorsal to the top.

Encouraged by the precise match in the early expression pattern we tested if the transgene provides enough protein to properly express *SxlPe*. *In situ* hybridization showed that *SxlPe* activity was indistinguishable from wild-type in *runt*^{3} mutants bearing the wild-type *runt·WRPY* transgene Fig.2.3A. and Fig.2.3B. Together the transgenic

runt expression pattern and the rescue of *SxlPe* activity suggest the *runt·WRPY* transgenes produce normal or near normal levels of *runt* protein.

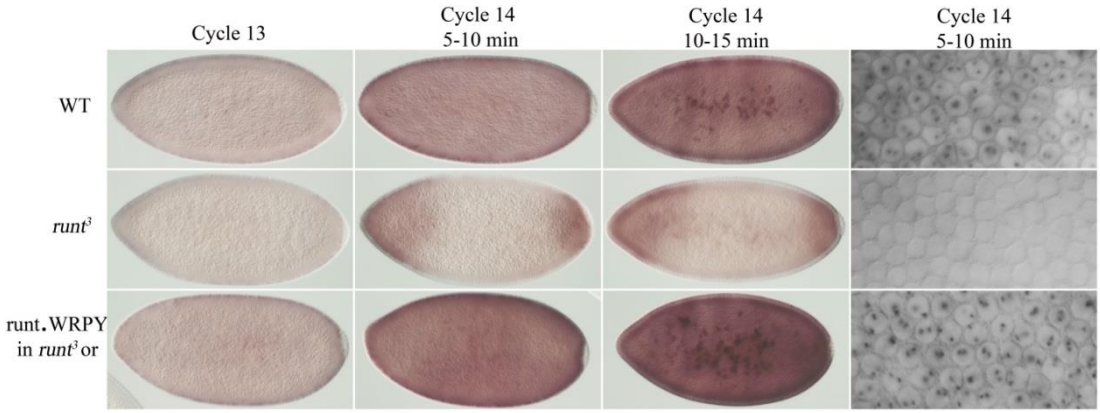

Fig.2.3. *runt·WRPY* transgenes provide full functions needed to express *SxlPe*. *in situ* hybridization to *SxlPe* transcripts in wild-type (WT) embryos, in *runt*^{3} mutants and in *runt·WRPY* bearing transgenic lines in *runt*^{3} mutants or heterozygotes. Because the *SxlPe* pattern is completely wild type in *runt*^{3} mutants bearing *runt·WRPY* transgenes, we cannot determine if the embryos shown in the bottom panel are *runt*^{3} mutants or *runt*^{3}/Balancer heterozygotes.

Loss of Runt's WRPY motif abolishes *SxlPe* expression affirm that the motif is essential for its activation function

To address the significance of Runt's Gro interactive motif, the WRPY motif was precisely deleted from the transgene to produce a *runt·*Δ*WRPY* transgenic line. Fig.2.4A. I used φC31 mediated integration to insert the Δ*WRPY* transgene in the same site in the genome where the wild type *runt·WRPY* transgene resides. I found Runt lacking its WRPY motif failed to rescue *SxlPe* expression in *runt*^{3} mutants. Indeed the *SxlPe* pattern

in *runt·ΔWRPY* bearing *runt*^{3} null mutants was indistinguishable from *runt*^{3} mutants alone suggesting that the Gro-interacting WRPY motif is required for Runt to activate *SxlPe*.Fig.2.4B.

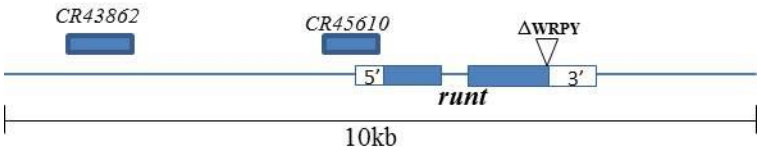

Fig.2.4A. *runt·ΔWRPY* transgene.

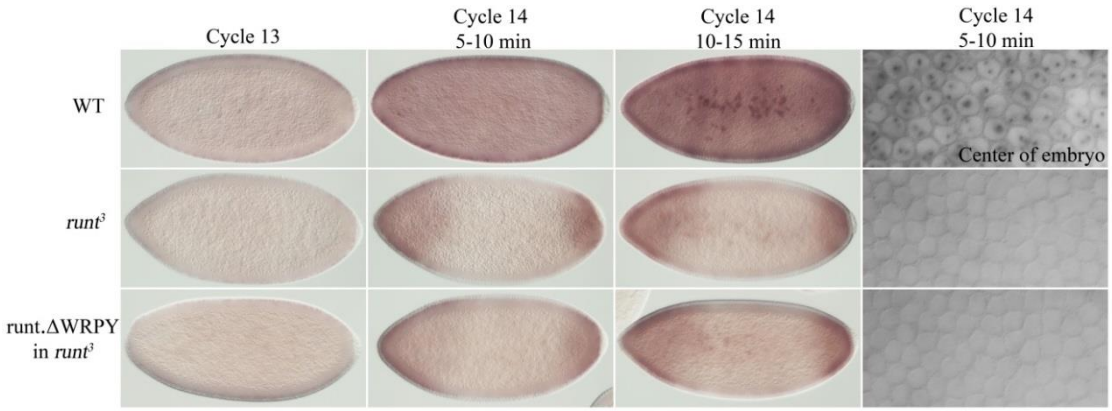

Fig.2.4B. *runt·ΔWRPY* transgenic lines lacking the C-terminal WRPY peptide fail to provide full functions needed to express *SxlPe* in *runt*^{3} mutant embryos. *in situ* hybridization to *SxlPe* transcripts in wild-type (WT), runt^{3} and in *runt*^{3} mutants bearing *runt·ΔWRPY* transgenes.

To ensure that the failure of the *runt·ΔWRPY* transgene to provide sex determination function was due to the loss of the WRPY motif, rather than from a lack of otherwise functional protein, I tested the *runt·ΔWRPY* transgene for its ability to

restore *ftz* expression in *runt*^{3} mutants. The ideal control for *runt·ΔWRPY* function would be to examine a target gene with an easily scored phenotype that is directly activated by *runt* independent of the WRPY motif, and that is expressed at the same time as *SxlPe*. No such gene exists but *ftz* is a reasonable choice because the *ftz* pattern is severely altered in *runt*^{3} mutants, and because the *ftz* pattern can be partially restored by *GAL-4* mediated overexpression of Runt protein lacking the WRPY motif [56].

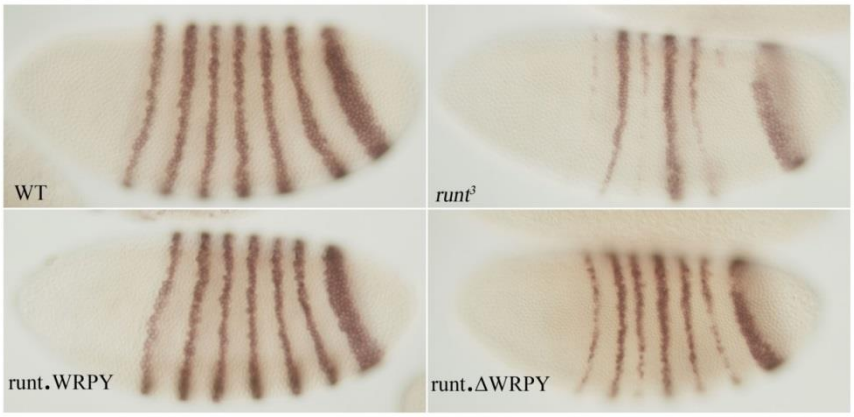

Fig.2.5. *ftz* expression in late nuclear cycle 14 detected by *in situ* hybridization. *runt* transgenes rescue (*runt·WRPY*) or partially rescue (*runt·ΔWRPY) ftz* expression in *runt*^{3} mutants.

ftz expression begins in early cycle 14 embryo and resolves into a clear seven stripe pair-rule pattern as cellularization progresses. Initially the first six stripes are expressed in 4 cells wide swaths whereas the seventh stripe is expressed more broadly. By the end of cellularization *ftz* expression is restricted to two cells widths in all seven stripes [66]. In the absence of *runt*, *ftz* expression is close to normal during early

cellularization, but the *ftz* is reduced and prematurely aborted just prior to gastrulation. Fig.2.5. Ectopically expressed Runt without the WRPY motif has been shown to activate *ftz* and partially rescues its normal expression pattern [56]. I found that my *runt·ΔWRPY* provided a similar rescue of the *ftz* pattern. Fig.2.5. Runt·ΔWRPY protein clearly activates *ftz* in most of the cells suggesting that the transgene produces a functional Runt·ΔWRPY protein. The mechanism by which Runt regulates *ftz* is not known and is likely at least partially indirect [67]. I found that expression of the wild type *runt* protein from our *runt·WRPY* transgenes was capable of restoring *ftz* expression to a wild type pattern consistent with the notion that *runt's* control of *ftz* occurs via more than one mechanism.

Runt's Gro interacting motif when modified into a more potent Gro-interacting motif 'WRPW' it restores Runt's activation function at *SxlPe*

Deletion of the WRPY motif eliminates Runt's ability to maintain *SxlPe.* If the loss of interaction with Gro is responsible for the loss of activation function it should be possible to restore Runt's activation function by substituting a different Gro interaction motif. A well-known and potent Gro-interacting domain is found in the dedicated repressor proteins of the *hairy-E(spl)* family. *hairy-E(spl)* interact with Gro through their conserved C-terminus 'WRPW' motif and recruit it to target gene promoters [56]. The molecular interactions of Gro with WRPY and WRPW peptides are similar except that the WRPW peptide interacts with considerably higher affinity. I constructed a *runt·WRPW* transgene by changing 'Y' into 'W' in or wild type transgene and inserting into the same region in the genome as other two transgenes we tested. Fig.2.6A.

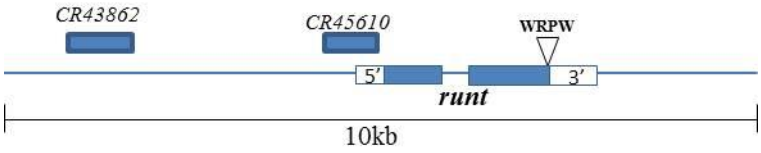

Fig.2.6A. *runt·WRPW* transgene.

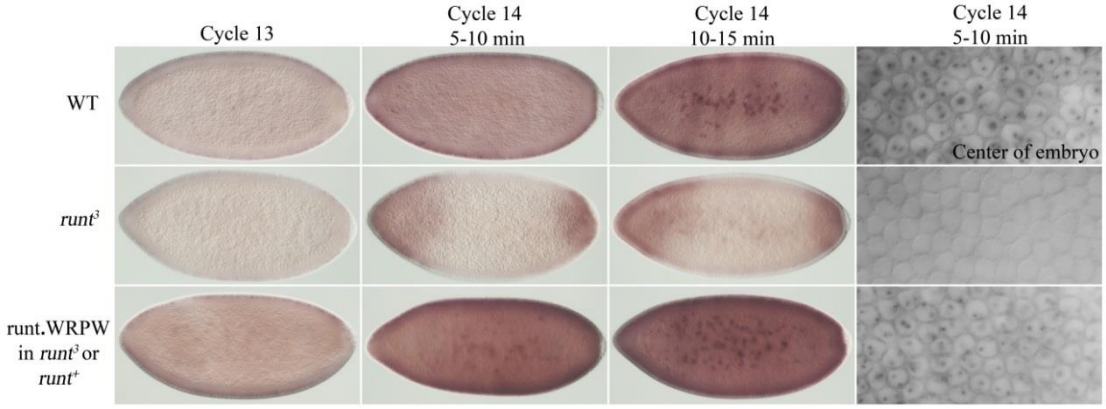

Fig.2.6B. *runt·WRPW* transgenes provide full functions needed to express *SxlPe* in *runt*^{3} mutants. *in situ* hybridization to *SxlPe* transcripts in wild type (WT), *runt*^{3} and *runt·WRPW* transgenic lines in *runt*^{3} or *runt*^{3}/ Balancer embryos. Because the *SxlPe* pattern is completely wild type in *runt*^{3} mutants bearing *runt·WRPW* transgenes, we cannot determine if the embryos shown in the bottom panel are *runt*^{3} mutants or *runt*^{3}/Balancer heterozygotes.

I was curious about the *runt·WRPW* transgenes as we wondered if the addition of WRPW might change Runt from an activator to a repressor as proteins with WRPW motif are dedicated repressors. Instead, we found that the Runt·WRPW protein was able to provide full function at *SxPe. In situ* hybridization to the *SxlPe* transcripts confirmed the expression restored to normal levels at the central broad region of female

embryo.Fig.2.6B. Runt thus acts as an activator at *Sxl* so long as its C terminus contains either a WRPY or WRPW Gro interaction motif.

How might Runt 'poison' Gro function? One possibility is that Runt might bind Gro throughout the embryo and prevent or limit its binding via a 'titration' type mechanism. I do not favor a general titration scheme for several reasons including the lack of *runt* dose-sensitivity, but the most compelling reason is that Runt appears also to require DNA binding to activate *SxlPe*. A requirement for Runt DNA binding in *Sxl* activation was first reported by Kramer *et al* [35]. They found that a mutant carrying two amino acid changes, C127S and K199A (CK), that were thought likely to disrupt DNA binding without greatly perturbing Runt structure, was unable to activate *SxlPe* when overexpressed. Because more recent work on Runt structure suggests that the CK mutation might actually have altered Runt protein structure, I reexamined the requirement for Runt DNA binding by creating a series of mutations that, based on structural studies, should inhibit DNA binding without disturbing overall protein structure [35] [68]. I examined mutations in the Runt domain that disrupt specific interactions with DNA, R80A that disrupts base and phosphate interactions, R139A that disrupts a phosphate contact, R142A that disrupt a base contact and the phosphate backbone outside the 6 conserved bases of Runt binding, R174A and R177A that disrupt specific base interactions. Fig.2.7. None of the mutant transgenes that altered Runt DNA binding were capable of rescuing *Sxl* expression in the $runt^3$ null mutant background confirming that Runt must be capable of binding DNA and of binding Gro in order to activate *SxlPe*.

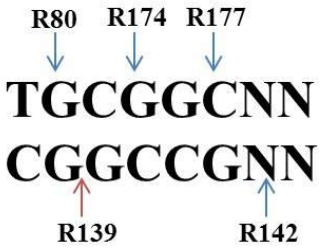

Fig.2.7. Modifications in 'Runt' domain that interact at different DNA positions.

Runt sequence specific versus non-sequence specific binding at *Sxl*

One impetus for examining the specific DNA binding mutants I created was to explore the possibility that the DNA binding requirement for Runt function may reflect a requirement for low-affinity or sequence non-specific binding. While it has been claimed Runt regulates *SxlPe* by sequence-specific DNA at the promoter there are no Runt binding consensus sequences at the promoter, nor even sites with one mismatch [35]. Moreover, the published evidence for sequence-specific binding is that Runt/CBF-beta binding to 200-300 bp fragments of *Sxl* promoter DNA could be competed off by a bona fide Runx binding site from the Polyoma virus enhancer. This is an inappropriate assay to assess sequence specific binding, as high affinity binding sites would effectively compete both weak sequence-specific and nonsequence-specific DNA binding. A requirement for non-specific DNA binding is entirely consistent with experiments performed by former PhD student Dun Yang in our lab, who documented only low affinity, non-specific interactions between Runt monomers, and Runt/Bro or Runt/Bgb dimers and the *Sxl* promoter. Unfortunately, my failure to recover any functional Runt

transgenes with DNA binding mutations did not allow me to draw solid conclusions about the need for sequence-specific versus sequence non-specific DNA binding.

As an alternative approach I considered the possibility that high-affinity binding of Runt might be associated with repression whereas lower affinity or sequence non-specific interactions might permit activation via antagonism of Gro function. I therefore introduced the conserved Runt DNA binding site, 'TGCGGT', into *SxlPe-lacZ* fusion transgene between -400 and -800 as this appears to be a spacer region with few binding sites or regulatory impact [43] [30]. Fig.2.8. I found that expression from promoters with on, two, or three added Runt binding were all indistinguishable from the control *SxPe-LacZ* transgene (Data not shown). Another transgene that had three tandem repeats of the Runt binding site was also wild type with no expression in males.

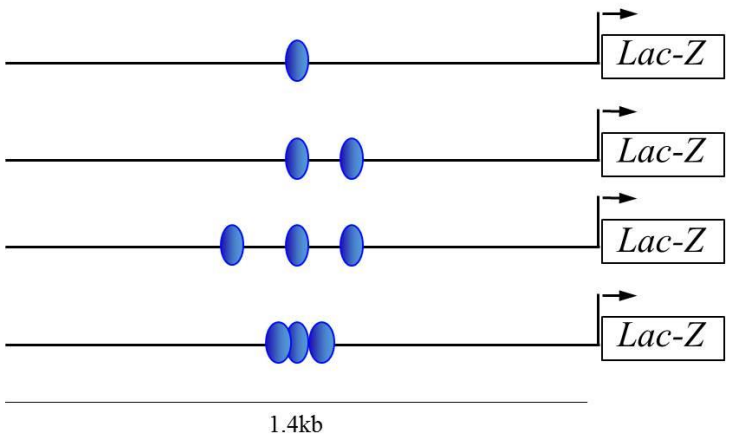

Fig.2.8. *SxlPe-lacZ* constructs with Runt binding sites added to the promoter. Blue ovals represents 'TGCGGC' site.

Chapter summary

The goal of the experiments described in this chapter was to understand the molecular mechanism of Runt's control of *SxlPe* activity. I found that Runt lacking its WRPY motif cannot function as an activator at *Sxl* suggesting that Runt must interact with Gro corepressor in order to activate *SxlPe*. My conclusion was strengthened by showing that the higher affinity WRPW Gro-interacting sequence from dedicated repressors also allowed Runt to activate *SxlPe*. I propose that the paradox posed by a requirement for a corepressor interaction in transcriptional activation is solved by Runt binding to and antagonizing Gro's potent repressive activity at *SxlPe*. Runt appears to be required to maintain *SxlPe* activity is not to initiate it. This suggests that the initial activation of *SxlPe* depends primarily on Scute/Da and SisA but that those factors alone are insufficient to maintain activity in the presence of Groucho. In this context, the two X female dose of Runt is seen as effectively dampening Gro function at an already active *SxlPe* so as to maintain its activity. Runt and other Runx proteins are known to activate and repress target genes in a context dependent fashion. At *SxlPe* that context includes a remarkable degree of functional specialization of the XSEs and other factors that regulate the promoter. I suggest that *SxlPe* is not unique in these respects and that Runt's ability to mediate activation and repression via multiple mechanisms make it likely that Runx proteins will be mediate many similarly complex developmental decisions.

CHAPTER III

ROLE OF *BIG BROTHER* AND *BROTHER* IN *SXL* REGULATION

CBF-β factors form heterodimers with Runx proteins. The Runx/CBF-β dimers bind DNA with much greater affinity than Runx proteins alone and dimerization is generally thought to be essential for mammalian Runx function [69] [70] [71] [72] [73]. There has, however, been considerable speculation, and at least one report, that some Runx functions may not require interactions with CBF-β [61]. The experiments presented in this chapter were directed at answering the questions of whether Runt's function at *Sxl* requires either of the two fly CBF-β proteins, Big brother (Bgb) and Brother (Bro). Published data from experiments using microinjected RNAs indicated that one or more of the CBF-β proteins was likely needed to activate *SxlPe*; however, ambiguities in those experiments and indications that sequence-specific DNA binding may not be needed for Runt's control of *Sxl*, suggested the possibility that *Sxl* might represent a CBF-β independent target for Runt proteins in flies. I begin with an overview of published and unpublished experiments relevant to Runt and *Sxl*, I then consider what is known about *Drosophila* CBF-β functions in other developmental contexts and conclude with descriptions of my experiments to isolate and study CBF-β single and double mutants.

CBF-β functions in *Drosophila* sex determination

Peter Gergen's laboratory was the first to address the role of CBF-β proteins at *Sxl*. Because there were no functional Runt transgenes and because of the relatively weak genetic effects of *runt* on *Sxl*, Gergen's lab monitored *SxlPe-lacZ* expression after microinjection of Runt or CBF-β RNAs [35] [59]. They found that injection of wild-type *runt* mRNA was able to rescue *SxlPe* activity in the central region of $runt^{LB5}$ null embryos, and that the runt mRNA induced ectopic *SxlPe-lacZ* expression in about half of male embryos. Li and Gergen found that CBF-β interactions were likely important for *SxlPe* activity because RNA encoding the *runt* mutant [G163R] that strongly decreased Runt/CBF-β interactions in two-hybrid assays was no longer capable of inducing ectopic *SxlPe-lacZ* expression. Curiously, injection of *runt* [G163R] mRNA also interfered with activation of *SxlPe* in female embryos. Li and Gergen interpreted this to mean that, when highly overexpressed, runt [G163R] protein was likely capable of binding to *SxlPe* and displacing wild-type Runt or interfering with wild-type Runt's action at *SxlPe*.

An important caveat in these experiments is that microinjection of mRNA produces significantly higher levels of protein than are expressed endogenously, or even by the maternally driven GAL-4/UAS system. None of the effects on *SxlPe* reported by Li and Gergen were observed when wild-type *runt* or the *runt* [G163R] mutant were expressed using maternally driven GAL-4 and UAS expression constructs [59]. Likewise, ectopic induction of *Sxl* in males is not observed when $runt^{+}$ is present in two or even three copies [23] (Mahadevarju unpubished).

Redundant or partially redundant CBF-β functions in *Drosophila*

In their initial characterization of *Bgb* and *Bro,* Golling and Gergen reported erroneously, that both genes were expressed maternally, meaning that both *Bgb* and *Bro* proteins should be present in the embryo at the time *SxlPe* is expressed [38]. Since both proteins work similarly in in vitro assays, the Gergen lab did not address the question of whether there were functional differences between the two CBF-β proteins.

The issue of CBF-β specificity was addressed more directly by Kaminker *et al*. who studied *Bro* and *Bgb* functions in the fly embryo, eye, and nervous system using both genetic and RNAi-based approaches [60]. They found that *Bgb* mutations interacted synergistically with mutations in the Runx homolog, *Lozenge,* to disrupt the structure of the fly eye. Homozygous clones of *Bgb* mutants, however, failed to show any eye defects suggesting that *Bro* can compensate for the loss of *Bgb* in the eye and thus that the two CBF-β proteins are normally fully redundant in eye tissues. In contrast, in embryonic segmentation, Kaminker *et al*. found evidence for partial redundancy. Their key observations were that knock-down of *Bgb* function was without effect on segmentation, whereas knock-down of *Bro* function caused strong segmentation defects similar to those seen with partial loss-of-function *runt* alleles. When *Bro* and *Bgb* were together targeted by RNAi much stronger segmentation defects were observed. This suggests that the two proteins play partially redundant roles in segmentation and hint that Runx proteins may prefer to interact with Bro when both CBF-βs are available. In the peripheral nervous system, however, there was evidence of a *Bgb*-specific effect. There, *Bgb* alleles caused defects similar to those observed in *runt* null mutants suggesting that

only *Bgb* can provide CBF-β function in the affected cells. Because Kaminker *et al*. did not monitor *Bgb* and *Bro* mRNA or protein levels, it is not known whether the full or partial specificities observed for *Bgb* and *Bro* in segmentation, and in the PNS, represent differences in CBF-beta expression or from functionally divergent amino acid sequences between Bgb and Bro.

Bgb and *Bro* are both present in the early embryo but are differentially expressed

The first published reports of *Bro* and *Bgb* expression found that both mRNAs were present in embryos but the two papers reached different conclusions as to their patterns of embryonic expression. Golling and Gergen reported that both were expressed maternally, but Fujioka *et al*., reported that while *Bgb* was provided maternally, *Bro* was expressed zygotically around the time of cellularization [38] [37]. My *in situ* hybridizations confirm the findings of Fujioka *et al.* that *Bgb* is expressed maternally, and that *Bro* is expressed only for a brief period prior to cellularization.Fig.3.1. Interestingly, the temporally dynamic early expression pattern of *Bro* during nuclear cycles 10-14 closely matches the patterns of the XSE genes, *sisA* and *scute*, raising the possibility that *Bro* might have a role in *Sxl* regulation that cannot be fully filled by maternally supplied *Bgb*. A possibility also supported by the apparent preference for *Bro* over *Bgb* in early embryonic segmentation [60].

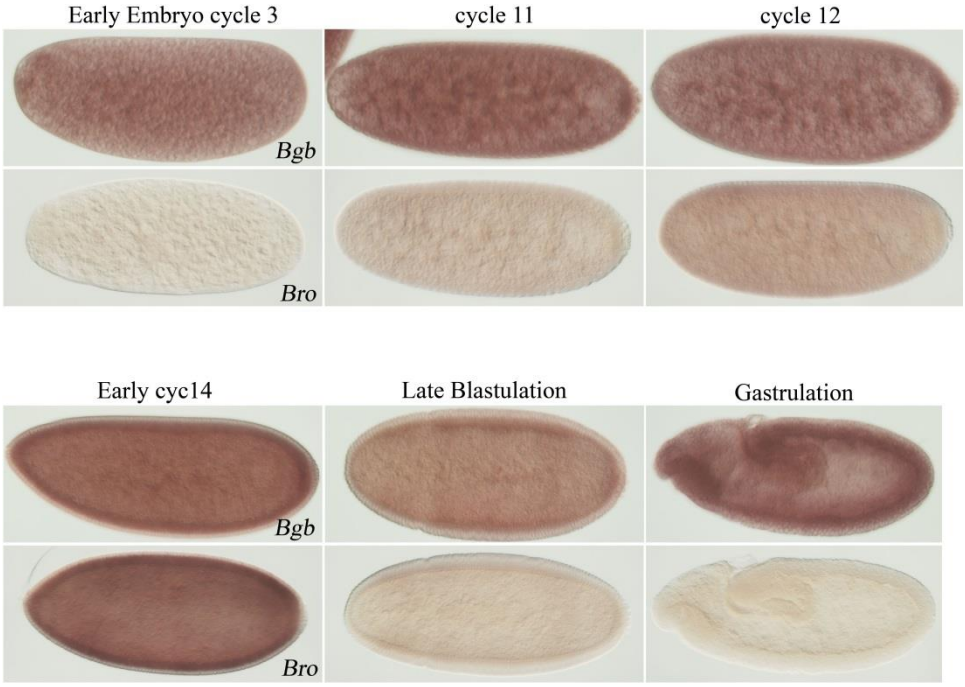

Fig.3.1. *In situ* hybridization to *Bgb and Bro* mRNA in WT embryos. *Bgb* mRNA (top panels) is maternally deposited in the eggs and expressed much later in zygote. *Bro* is expressed similarly to XSEs. Bro transcription extends from cycle 11 to early 14 but declines rapidly thereafter.

Maternal *Bgb* appears not to be needed for *SxlPe* activity

Big brother is an essential gene as mutants carrying the strong loss-of-function allele $Bgb^{KG03779}$ are not viable. To determine if maternally supplied *Bgb* is needed for *Sxl* expression, Elena Khozina, a former Master's Degree student in my lab, generated germline $Bgb^{KG03779}$ clones using the dominant female sterile technique (DFS) to eliminate maternally supplied *Bgb* protein and examined the progeny for Sxl_{pe} activity. Her finding, that *SxlPe* was activated and expressed normally in the absence of maternal *Bgb* function (data not shown), is consistent with three possible explanations. First, the

two fly CBF-beta proteins may function redundantly at *SxlPe.* Second, that zygotically expressed *Bro* exclusively provides the CBF-β function needed for Runt to regulate *SxlPe*, and third, that neither CBF-β protein is required for Runt to properly regulate *Sxl*.

Brother is not needed for _SxlPe_ expression nor is it required for viability or fertility of _Drosophila_

Since no clear loss-of-function *Bro* mutants were available, I constructed a *Bro* null mutation at the endogenous locus using gene targeted ends-out site directed mutagenesis [74] [75]. Fig.2.2. I deleted the first 135 amino acids (of 213 total), which include all the residues known to be required for Runx protein binding**,** to ensure complete loss of function [68]. It replaced the deleted amino acids with a w^+ gene cassette flanked by *lox* recombination sites to serve as a genetic marker.

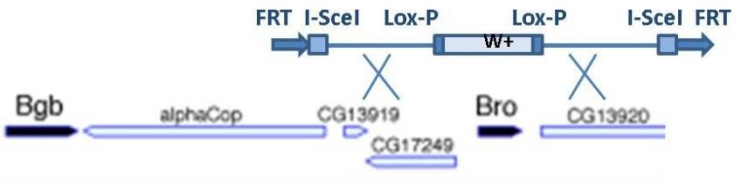

Fig.3.2. Generation of a _Bro_ deletion mutant by 'ends-out' homologous recombination. A donor chromosome generates a circular fragment with *w+* cassette with homology to either side of the *Bro* gene was linearized by I-SceI in the germ line. Recombination between the linearized fragments and the chromosome resulted in the replacement of *Bro* with the *w+* marker and *Bro* deletion.

If *Bro* alone is involved in *Sxl* regulation, as was hinted by the *Bro* expression pattern (Fig.1), then the *Bro* homozygous mutant phenotype should mimic that of a *runt* null allele and eliminate *SxlPe* expression in the central region of female embryos. However, the *ΔBro*^{1} deletion allele had no observable effect on *SxlPe* expression.Fig.3.3. This indicates either that the CBF-β proteins functions are indeed redundant at this early stage, or that both CBF-β are dispensable for Runt's regulation of *SxlPe*.

Fig.3.3. *In situ* hybridization to *Sxl* mRNA in Δ*Bro*^{1} embryos. *Sxl* mRNA is expressed similar to WT, begin to express from cycle 12 to early 14 and decline later.

Surprisingly, given the previous reports of a significant requirement for *Bro* function in early embryonic segmentation, homozygous *ΔBro*^{1} flies were completely viable, fully fertile, and normal in appearance [60]. This is in contrast to *Bgb* mutants which are lethal to the organism [60] (Mahadevaraju and Erickson JW unpublished). These findings suggest that *Bgb* is expressed broadly enough to provide all needed CBF-β functions for the fly, and indicate that flies, like vertebrates can function with only a single CBF-β. The previously reported function for *Bro* in embryonic segmentation, thus, likely represents an experimental artifact of injection of double stranded RNA into

the embryo. Given, the sufficiency of Big brother, the question of why there are two CBF-β genes in flies remains unanswered.

Could *SxlPe* be expressed independent of the CBF-β proteins?

Big brother is essential for fly viability and the single CBF- β present in vertebrates is generally thought to be essential for RUNX proteins in mammals [53]. A recent study, however, reported evidence that mouse Runx1 may have functions that are independent of CBF-β [61]. Given that the primary function of CBF-β proteins appears to be in stabilizing high-affinity DNA binding of Runx proteins, we wondered whether *SxlPe*, which lacks detectable Runt binding sites, and appears not to be bound in a sequence specific manner by Runt/ CBF-β heterodimers in vitro, (see Introduction, Chapter I), might represent a case where Runt acts independent of CBF-β function.

Construction of a *Bgb Bro* double mutant

To determine if Runt's function at *SxlPe* is dependent or independent of the CBF-beta proteins, I needed to construct and analyze *Bgb Bro* double mutants. While simple in principle, the extremely close linkage of *Bro* and *Bgb* posed several problems for the construction of the needed double mutant. As described in the following paragraphs, I ultimately tried four different approaches before isolating the necessary *Bgb Bro* double mutant.

An attempt to construct Bro Bgb double mutants by male recombination

The *Bgb* and *Bro* genes are closely linked, separated by only 7.2 kb. This close linkage precludes the use of normal homologous recombination to generate a double mutant. Instead, I attempted to use *P*-element induced male recombination to create a $Bgb^{KG03779}$ ΔBro^{1} double mutant. In *P*-mediated genetic exchange, recombination is induced in males by providing *P*-element transposase, which causes breakage and recombination at the site of integrated *P*-elements [76]. The $Bgb^{KG03779}$ mutation is caused by a *P*-element insertion, carrying the w^{+} and y^{+} markers, into the *Bgb* locus and a dominant marker *Pr* is present on the other arm of the chromosome. I screened for males with w^{+} and y^{+} markers and loss of *Pr* indicating the recombinants.Fig.3.4.

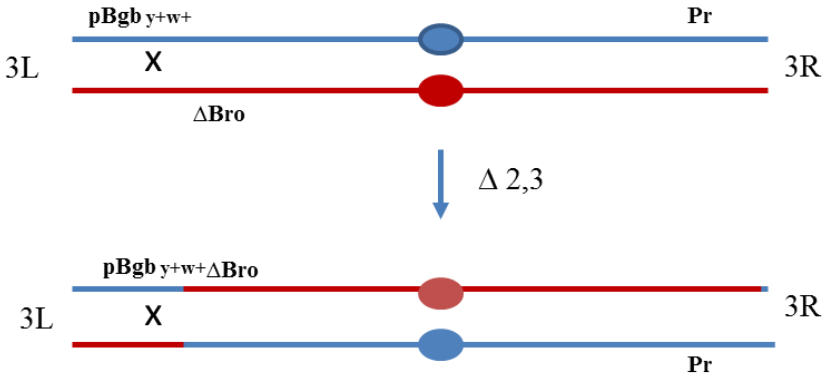

Fig. 3.4. Plan to create Bgb^{P} ΔBro^{1} double mutants by *P*-element induced male recombination. Males with Bgb^{P} and ΔBro^{1} chromosomes that produce transposase Δ2, 3 induce recombination at the proximity of *P* element. Putative Recombinants were selected as y+ w+ files that lost the dominant bristle marker *Pr*.

49 flies were selected as candidates to be double mutant by genetic markers. They were checked to see if they would fail to complement the $Bgb^{KG03779,}$ the strong *Bgb* loss of function allele. Most of the candidates were fully viable over $Bgb^{KG03779}$ and the others exhibited only slightly reduced viability. All 49 candidates were viable as homozygotes. It was strange to recover viable flies that genetically indicate the presence of $Bgb^{KG03779}$; however, male recombination is known to cause abnormalities, such as duplications and other rearrangements, and it may be that such events generated complex recombinants that contained both wild type and mutant alleles. Since I was unable to recover double mutants by male recombination, I abandoned the approach in favor of the 'ends-out' targeting method I used to generate the ΔBro^{1} allele.

<u>An attempt to generate *Bgb Bro* double mutant by homologous recombination</u>

My approach to generate a deletion of *Bgb* was similar to what I used for *Bro*. Fig.3.5. The injection plasmid was transformed into a w^{-} derivative of ΔBro^{1} allele (which I had constructed so that the w^{+} marker could be excised using Cre-lox recombination). I carried out the standard ends out targeting protocol. In addition, I used a similar strategy to delete *Bgb* and *Bro* along with the genes in between them using a fly with a Bgb^{+} Bro^{+} third chromosome as the target. We failed to recover any double mutants using either approach. Reasoning this unexpected result, and perhaps the failure of P-recombination, may have resulted from chromosomal rearrangements around *Bgb*. I checked the structure of the *Bgb-Bro* region by PCR amplification but found no evidence of any abnormalities. All genes were present and in the proper orientation.

We, thus, do not have an explanation for why P-induced recombination and targeted knockout failed to isolate *Bgb Bro* double mutants.

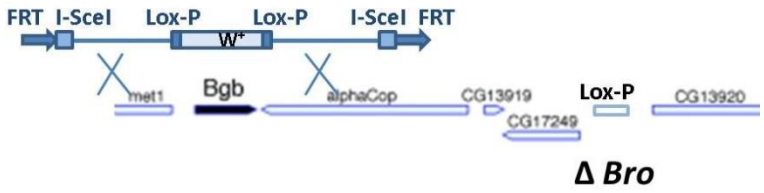

Fig.3.5. Plan to generate *Bgb* deletion mutant by 'Ends-out' homologous recombination in Δ*Bro*[1].

An EMS mutagenesis screen to isolate *Bgb* Δ*Bro* double mutants

Because *Bgb* appears to be an essential gene we reasoned that it should be possible to isolate an EMS-induced *Bgb* allele by looking for mutations that failed to complement the $Bgb^{KG03779}$ allele. Using the ΔBro^{1} genetic background as the starting point we screened the progeny of 5,000 single F2 males to find possible double mutants. I recovered 6 candidates that were lethal in combination with $Bgb^{KG03779}$ however, none of the 6 showed sequence change in *Bgb*, and all complemented a newly available deficiency allele, *Df(3L) BSC870,* that removes *Bgb* and several adjacent genes, but leaves *Bro* intact. This suggests that I isolated mutations that failed to complement another unidentified lethal mutation on the $Bgb^{KG03779}$ chromosome. While I could have employed a larger scale and improved EMS screen and improved the screen by removing the extraneous lethal mutations from the *Bgb* mutant, we decided to switch

approaches yet again to make use of the newly developed CRISPR technology as it promised a more directed and faster means to isolate the double mutant.

CRISPR targeting to isolate a *Bgb* Δ*Bro* double mutant

CRISPR (Clustered regulatory interspaced short palindromic repeat) is a bacterial three part RNA guided defense system against viruses. Bacterial CRISPR RNA (crRNA) specifically binds to viral DNA, transactivating CRISPR RNA (tracrRNA), recruits Cas9 endonuclease to the complex leading to a double stranded DNA break in the target genome. This system has been simplified and adapted for use in *Drosophila* and several model organisms, including humans, for genome editing. This simplified system uses a single targeting RNA, called an sgRNA, which fuses crRNA and the tracrRNA sequences, and a source of Cas9 enzyme. Target specificity is defined by a 20 bp sequence in the crRNA portion of the sgRNA that is homologous to the target. Cas9 enzyme then cleaves the target and non-homologous end joining (NHEJ) generates small insertions or deletions [77] [78] .

I exploited a Crisper system in which plasmids expressing the sgRNA from the U6b promoter are injected into embryos expressing Cas9 in their germ cells. *Bgb* was targeted by CRISPR in ΔBro^{1} flies. Guide sequences for targeting were selected through a searchable database that predicts off targets [79].

I recovered a single base deletion in *Bgb* that shifts the *Bgb* reading frame. Deletion of the C base at +182 changes the 61st amino from tyr to ala [Y61A] and all subsequent amino acids up to the introduced stop codon. The sequence change suggested

that the new *Bgb*^{8} allele likely produces a nonfunctional truncated *Bgb* protein. I confirmed this by showing that the *Bgb*^{8} *ΔBro*^{1} combination fails to complement the strong *Bgb*^{KG03779} allele and is homozygous lethal.

As a backup strategy I simultaneously targeted both *Bgb* and *Bro* in wild-type flies using guide RNAs for both genes. In choosing the sequences for the guide RNAs we failed to notice that they shared a sequence segment that targeted homologous regions of the two genes. Consequently, I recovered a Crispr-induced allele that deleted the 8 kb region between *Bgb* and *Bro* and created a precise in frame fusion of the proximal half of *Bgb* (61amino acids) and the distal half of *Bro* (143 amino acids). This fusion mutation, which I refer to as *Bgb-Bro*^{6} failed to complement *Bgb*^{KG03779} suggesting the fusion- protein cannot simply substitute for wild type *Bgb*. Since I do not know which sequences control the expression of *Bgb* or *Bro* we do not know whether the failure of the fusion protein to complement reflects altered regulation or altered function of the fusion protein.

Germ line clones of *Bgb*^{8} Δ*Bro*^{1} double mutants are inviable

I attempted to generate germ line clones of *Bgb*^{8} Δ*Bro*^{1} and *Bgb-Bro*^{6} using the DFS technique but failed to recover any eggs from either of the mutants. This result was both disappointing and unexpected. Disappointing because the failure to isolate viable germline clones means that I cannot genetically address the question of whether there is a requirement for CBF-β function at *SxlPe*. Unexpected because previous work indicated that *Bgb*^{KG03779} mutants produced viable germline clones and viable offspring

(see above), and because *Bro*^{1} deletions are fully viable and fertile. The simplest explanation for the failure of *Bgb*^{8} Δ*Bro*^{1} germline clones to produce eggs is that *Bgb* and *Bro* can function redundantly in the female germline. In the "redundancy scenario", the viability of *Bgb*^{KG03779} clones would be explained by the redundant *Bro*$^{+}$ gene providing the needed CBF-β function. Likewise the viability of Δ*Bro*^{1} germlines would be explained by redundant *Bgb*+ function. A second explanation for the inviability of *Bgb*^{8} Δ*Bro*^{1} germ line clones is that *Bgb* alone may be required in the germline, a result possible if the strong *Bgb*^{KG03779} allele retains sufficient *Bgb* function to rescue the germline defect. A third alternative is that our lab's earlier experiments with *Bgb*^{KG03779} were incorrect and that our experimental approach was guided by false findings. A fourth, but unlikely, alternative is that a recessive germline lethal mutation was unexpectedly created in the course of our Crispr mutagenesis.

Is there another means to examine the function of CBF-β proteins at *SxlPe*?

Because of the importance to our lab's earlier finding that *Bgb*^{KG03779} germline clones were viable, I recently repeated those experiments and successfully generated both *Bgb*^{KG03779} clones, and viable progeny from the mutant germlines. This suggests the two most likely explanations for our failure to obtain viable *Bgb Bro* germline clones is either the *Bro* can provide *Bgb* function in the germline or that the *Bgb*^{KG0377} allele retains partial function. Either scenario offers hope that it may still be still possible to address the functional role of the CBF-β proteins at *SxlPe*. If *Bro* can rescue the germline defect of *Bgb* mutants, it should be possible to provide that function from a *Bro*$^{+}$ transgene. In this case, we could induce *Bgb*^{8} Δ*Bro*^{1} germline clones in flies

carrying an unlinked *Bro*$^{+}$ transgene. Since *Bro* mRNA is not normally found in the embryos, it may be that the transgenic copy of *Bro*$^{+}$ would complement the CBF-beta germline defect, and still produce embryos lacking both *Bro* and *Bgb* function. If instead, *Bgb*^{KG0377} retains partial function, I can use Crispr to induce a *Bro* mutation on the *Bgb*^{KG0377} chromosome and examine *Sxl* expression in embryos where the only source of CBF-β protein would be the defective maternal *Bgb*^{KG0377} alleles. (The fourth alternative, that an unknown germline sterile mutation was induced during Crispr, could be addressed in a number of ways, most simply by using Crispr to isolate several new Bgb Bro double mutants).

Flies survive with one copy of *Bgb*. Initial genetic analysis of *Bgb*^{8} Δ*Bro*^{1} allele

Although I could not analyze *Bgb*^{8} Δ*Bro*^{1} germline clones I did undertake an initial genetic analysis to determine if our new *Bgb* and *Bro* mutants exhibited any obvious abnormalities or genetic synergisms characteristic of many genes that regulate *Sxl*. Flies with only one copy of *Bgb*$^{+}$ (*Bgb*^{8} Δ*Bro*^{1} /*Bro*$^{+}$ Δ*Bro*^{1}) are viable and fertile and have no visible defects. This was surprising as there are normally four copies of CBF-beta factors in flies. With respect to sex determination, the *Bgb*^{8} Δ*Bro*^{1} alleles showed no lethal interactions with the XSE genes *sisA*, *sisB* or *runt*. The lack of any synergism with the XSE genes is difficult to interpret, except to say that if *Bgb* is

involved in regulating *Sxl*, one maternal copy is more than sufficient to provide the necessary CBF-β function.

Chapter summary

The experiments conducted in this chapter were to answer whether Runt's function at *Sxl* requires either of the two fly CBF-β, Brother (Bro) and Big brother (Bgb). A ΔBro^{1} null mutant constructed by ends out targeting showed no effect on *Sxl* and flies are fully viable and fertile. It provides new information that flies, like mammals can function with only a single CBF-β gene, and shows that some previous inferences about the functional specificity of the fly CBF-β proteins were incorrect [60]. Unfortunately, I have been unable to address the critical question of whether CBF-beta function is required or dispensable at *SxlPe* because we have been unable to generate embryos lacking both proteins. While deeply disappointing there is hope that one of two revised strategies will allow the question of which if any CBF-β proteins function with Runt to regulate *SxlPe*.

CHAPTER IV

CONCLUSIONS AND METHODS

Numerous fundamental developmental decisions are taken in response to small quantitative differences in signal molecules. Often these signals are rendered for brief periods during early development and lead to distinct and permanent cell faces. Sex determination in *Drosophila* is a particularly illustrative case of where a developmental fate decision is made in response to a small and clearly defined difference in doses of signaling elements. A twofold difference in the concentrations of four X-linked XSEs *sisA*, *scute*, *upd* and *runt* is sufficient to define opposite sexual fates. Providing twice as much XSEs in female embryos compared to males stably activates the *SxlPe* promoter and initiates the cascade that produces the *Sxl* protein needed for female development. We, and others, have proposed this two dose difference in the XSE-based signal must be amplified in some way to serve as a reliable indicator of sex. Previous work from our lab suggested that signal amplification primarily depends on Groucho (Gro), a member of the Gro/TLE1 family of co-repressors [44]. Gro is maternally contributed and recruited to the *SxlPe* promoter, most likely via several DNA binding protein including the hairy/E(spl) (Hes) related protein Deadpan (Dpn) [63].

In the absence of Gro, *SxlPe* is expressed in both the sexes in direct proportion to X chromosome dose. *SxlPe* is also expressed earlier than normal in XX embryos lacking Gro. This suggests that Gro sets a kind of repression threshold that is normally overcome in females because of the 2X doses of XSEs. Gro also ensures that *SxlPe* never comes

on in males and we proposed that this occurs in part by males actively modulating the threshold upwards during the period in which sex is assessed at *SxlPe*. We have hypothesized that antagonism of Gro function in females may be a key part of the signal amplification mechanism [44]. In this thesis I set out to test the hypothesis that the function of the XSE protein Runt is primarily to dampen Gro-dependent repression of *SxlPe* [63].

Runt is needed only in the regions where Gro is fully active

The notion that Runt might act as an activator of *SxlPe* by antagonizing Gro repressor function is not new. It was originally proposed in the paper that first showed the direct physical interactions between Runt and Gro, but the notion fell by the wayside when a subsequent paper made the claim that Runt was a direct activator of *SxlPe* [56] [35]. Several features of Runt's actions at *Sxl*, as well as a consideration of the mechanisms by which Gro regulates other cellular targets, have led us to return to the notion that Runt may serve as an activator of *SxlPe* by counteracting Gro-mediated repression.

The most important information was the correlation between the region specific effects of Runt on *SxlPe* and region specific regulation of Gro function by Torso RTK-dependent phosphorylation. In precellular embryos Gro is phosphorylated directly by MAPK at the embryonic poles [48]. The phosphorylation reduces the ability of Gro to repress target genes. Suggestively, the anterior and posterior regions where Gro is phosphorylated correspond to the regions where *SxlPe* activity does not depend on *runt*.

This raised the possibility that Runt is needed only in regions where Gro is fully active. This conjecture is nicely supported by early experiments showing that ubiquitous activation of Torso that leads to ubiquitous phosphorylation of Gro bypasses the requirement for runt in *Sxl* expression [23] supporting the requirement of Runt where Gro is active. As Gro is maternally contributed, but recruited to *SxlPe* by zygotically expressed Dpn at cycle 13 and 14 it is significant to know when Runt functions to antagonize Gro repression.

Runt is needed to maintain *SxlPe* activity, not to initiate it

Additional suggestive evidence that Runt antagonizes Gro came from careful analysis of *runt's* effects on the timing of *SxlPe* activity (Chapter II). *In situ* hybridization to nascent *SxlPe* transcripts revealed that *SxlPe* is activated normally at nuclear cycle 12 in the *runt* null embryos and that *runt* has only a small effect on *SxlPe* in the early stages of cycle 13. With the knowledge that the effect is huge in cycle 14, it suggests that Runt is not required for the initial activation of *SxlPe*, but is instead needed to maintain *SxlPe* in an active state during cycles 13 and 14. This is a very different behavior than exhibited by the XSE activators SisA and Sc which are required initially to activate *SxlPe*. It is consistent with the notion that Runt may function to help *SxlPe* remain active despite the continuing presence of Gro protein that is recruited through zygotically expressed Dpn during this time period.

Runt's WRPY motif is essential for its activation function

Runt interacts directly with Gro through its C-terminal conserved tetrapeptide motif, 'WRPY' [56]. Runt and other Runx proteins are known to recruit Gro to its targets via this WRPY motif to repress transcription [56]. Making the assumption that any antagonism between Runt and Gro would likely be mediated by the WRPY interaction motif, I tested whether *runt's* ability to function as an activator of *SxlPe* required the Gro interaction motif. I first showed that transgenes containing wild type *runt, runt·WRPY* expressed *runt* mRNA and controlled *SxlPe* in a manner indistinguishable from the endogenous *runt* locus. I then deleted the WRPY sequence from the transgene and found that the Runt·ΔWRPY protein failed to provide *runt* function at *SxlPe*. As a control I showed that the same *runt·ΔWRPY* transgene was able to partially rescue expression of *ftz*, a target gene previously reported to be partially dependent on the WRPY motif. I then asked if I could restore *runt* function by adding back a related tetrapeptide motif, WRPW, that mediates Gro-dependent repression in the Hes group of proteins, all of which are dedicated repressors [63]. The Runt·WRPW protein provided full activation and restored to normal levels of *Sxl* mRNA in the central broad region of female embryos. Together, my experiments indicate that for Runt to activate *SxlPe*, it must carry a functional Gro-interaction motif. The simplest interpretation is that Runt's functions as an activator by binding and inhibiting Gro's co-repressor functions at *SxlPe*. I will speculate on how Runt might accomplish this after I discuss the genetic and mechanistic implications of having XSE elements that function to maintain, but not initiate *SxlPe* activity.

Maintenance versus initiation functions for XSEs

My observations suggest that Runt is distinct from the strong and highly dose-sensitive XSE activators *SisA* and *Sc* because it is not required for initial activation of *SxlPe* and because it is only required in certain regions of the embryo. Curiously, Runt is not unique in these respects, as the XSE gene *upd*, and its maternally supplied targets in the JAK/STAT pathway, are also dispensable for transcription initiation at *SxlPe*.

Upd is a ligand for JAK/STAT pathway that ultimately helps direct activated STAT transcription factor to *SxlPe* [65]. STAT has been suggested to maintain the active chromatin by interacting with components of Brahma complex, a suggestion consistent with STAT brahma interactions in the induction of mammalian cytokine genes [65] [80] [81] [82] [83].

This suggests a multi-step mechanism for *SxlPe* activation. One step is initiation primarily in response to the dose sensitive activators *Scute* and *SisA*. A second step is needed to maintain *SxlPe* activity, but this may involve two different mechanisms: reinforcement of active chromatin by Upd/STAT, and inhibition of Gro and repressive chromatin by Runt, to ensure foolproof regulation of *Sxl*. According to this notion, *SxlPe* will not be stably expressed unless Sc/Da and SisA initiate expression and Runt and the JAK/STAT pathway reinforce that activation. In males, any increases in the doses of the maintenance factors are rendered moot because of the failure of the single X levels of Sc and SisA to activate *SxlPe*. This sequential multistep mechanism where Runt is functioning at the reinforcement step might also explain *runt* having only

relatively weak female-lethal genetic interactions with *sisA* and *sisB* and in having virtually no effect on male viability similar to *upd*.

Runt a significant factor in X-signal amplification

A twofold difference in the concentrations of four X-linked XSEs, providing twice as much XSE protein in female embryos compared to males, is sufficient to stably express *Sxl*. As proposed, this two fold difference in the XSE signal must be amplified in some way to serve as a reliable signal. My finding and analysis fits well with what our lab has proposed earlier, a model for X-signal amplification but adds important new details[44][84].Fig.4.1.

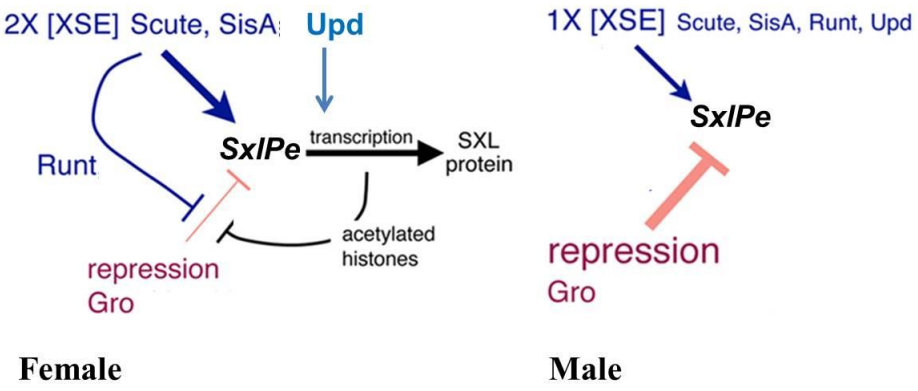

Fig.4.1. X-signal amplification model

The central feature of the modified X-signal amplification model is Gro and associate repressors that set an initial activation threshold. The threshold only overcome

when a sufficient amount of the strong activators Sc and SisA are produced in females to initiate *SxlPe* activation. Once transcription is initiated it is maintained in response to increasing levels of Sc and SisA that occur as a consequence of continued expression and nuclear multiplication, the effect of Runt antagonizing Gro to inhibit repressive chromatin, and Upd/Stat working to maintain active chromatin. In males, the 1X dose of *sisA* and *sisB* provides too little SisA and Sc to activate *SxlPe* initially. To prevent later activation in males, due to nuclear multiplication increasing the number of *sisA* and *sisB* templates in the embryo, the repression threshold is continually adjusted upwards in males by increasing the concentration of Gro via continued translation and increasing the recruitment of Gro to *SxlPe* via Dpn protein binding.

How might Runt antagonize Gro?

Groucho regulates target genes in part via chromatin modification but the molecular details of how Gro modifies the chromatin of its target genes remain unknown [85]. Despite the lack of detailed information about how Gro regulates chromatin, it is clear that phosphorylation is a key regulatory mechanism in dampening Gro function in flies and in other organisms. Homeodomain interacting protein kinase (Hipk2) promotes eye development via phosphorylating and inhibiting Gro to promote Notch signal transduction [86]. Gro phosphorylation is required to inhibition of cortical neuron differentiation. Hipk2 phosphorylates Runx1 in mammals and its associated factor p300 suggesting a possibility that Runt can recruit a kinase to phosphorylate Gro [87]. Most relevant here is that Gro is phosphorylated at the poles of the fly embryo to increase the

expression of genes needed to define the embryonic poles. The same phosphorylation event via torso kinase signaling likely affects Gro's function in sex determination as *tors*o gain of function alleles express *Sxl* uniformly in embryo lacking *runt* function [23].

While the Gro phosphorylation model for *Sxl* regulation is appealing, it has not been fully tested. Accordingly, I have developed several tools and outline plans that can be used to more rigorously test the idea that Gro phosphorylation is a central feature of *Sxl* regulation. The first approach is to perform the reciprocal experiments to those done by Duffy and Gergen, by asking if torso loss of function mutations compromise *Sxl* expression in the poles of the embryo [23]. This should lead to substantial reduction of *Sxl* expression in *runt; torso* double mutants since Groucho will effectively repress *Sxl* throughout the embryo. To directly test the role of specific Gro phosphorylations I have constructed Gro variants S308A, T510A (Gro^{AA}) and S308D, T510D (Gro^{DD}) that express non-phosphorylable and phosphomimetic Gro variants under the control of maternally driven Gal4. Gro^{AA} is expected to act as effective repressor of *SxlPe* whereas as the phosphormimetic Gro^{DD} should significantly reduce Gro ability to repress *SxlPe* throughout the embryo. Gro's functional status corresponding to its phosphorylation status in different mutant background can be analyzed by immunostaining as several antibodies are available to detect specific Gro phosphorylated forms.

I have also generated a eGFP tagged Runt transgene so that it will be possible to pull down Runt associated factors and identify them by mass spec analysis. With new techniques it is possible to isolate early embryo at a precise developmental time point when Runt is regulating *Sxlpe* [88] .

Sequence-specific DNA binding versus non-sequence specific binding of Runt at *SxlPe*

The available evidence suggests that Runt must bind DNA in order to activate *SxlPe*. The strongest evidence for this comes from my analysis of *runt* transgenes carrying 6 different mutations that should affect DNA binding without significantly altering Runt structure. In each case, the DNA-binding defective lines failed to provide the *runt* function needed to activate *Sxl*. One possibility is that Runt and a CBF-β partner bind to specific binding sites at *SxlPe*. Arguing against this are the lack of identifiable Runt binding sites near *SxlPe*, and the failure to identify sequence-specific DNA binding in vitro, using gel-shifts or DNase I footprinting [35] (D.Yang unpublished data). This raises the possibility that Runt activates *SxlPe* by binding DNA, but that it does so independent of specific DNA binding sites. We have proposed that Runt may interact with *SxlPe* by binding to other factors that regulate the promoter or by binding to active chromatin created by Sc/Da and SisA binding. In such a scenario, it might well be the case the Runt associates with DNA in a sequence non-specific manner. We have addressed this question in several ways but have been unable to determine whether specific or non-specific interactions are involved. One approach, discussed above, was to try to design Runt mutants that might be defective in sequence-specific binding but retain non-specific binding activity. All such candidates were defective for *runt* function. While this could indicate a requirement for sequence-specific interactions, it is impossible to rule out the possibility that the changes we introduced

affect all aspects of Runt's DNA binding. A second approach was to introduce specific consensus Runt binding sites into the *SxlPe* region. Our expectations for this experiment were quite open but we considered two general outcomes: First, that specific Runt binding might turn Runt into a repressor that reduced *Sxl* transcription. Second, that introduction of specific binding sites might lead to ectopic expression of *Sxl* in males by increasing the amount of XSE activators bound. Unfortunately, introduction of Runt binding sites did not detectably alter *SxlPe* activity leaving us with no means of interpreting our observations. Our final approach to assessing the possibility that Runt interact sequence non-specifically, was to analyze mutants defective for the fly's two CBF-β proteins, Bro and Bgb.

Roles of Bgb and Bro, Runt's putative partner in *Sxl* regulation

Core binding factors (CBF) consist of two subunits. The Runx subunit (CBF-alpha) binds DNA and is thought to mediate all transcriptional effects, whereas the CBF-β subunit acts to increase the DNA binding affinity of the dimer compared to a Runx monomer. Part of our rationale for this analysis was the idea that if Runt regulates *SxlPe* independent of high affinity binding sites, the CBF-β component might be dispensable for function. There is an apparent precedent for such a CBF-β-independent function in mammals; however, the paper making this claim does not provide definitive proof that this is a normal function of Runx proteins [61]. It is generally considered that the two fly CBF-β proteins are partially redundant based on genetic analysis and on the interchangeability of Bro and Bgb in *in vitro* binding assays [35] [60] (D.Yang

unpublished.) Several years ago our lab found, and I recently verified, that a loss of function allele that should eliminate or at least produce a defective maternal supply of Bgb, was without effect on *SxlPe* activity. Considering that redundancy between maternally supplied Bgb and zygotically expressed Bro, I created a deletion allele of *Bro* using site-specific mutagenesis. I found that ΔBro^{1} null mutants caused no detectable decrease in *Sxl* expression but also, surprisingly given the published literature, that ΔBro^{1} null mutants were fully viable and fertile and exhibited no other observable phenotypes [60]. While this result shows that flies, like mammals, can function with only a single CBF-β gene, it did not rigorously address the question of redundancy for *Sxl* regulation.

To determine if one or neither of the CBF-β proteins are needed for *SxlPe* activity I needed to analyze *Sxlpe* expression in germline clones of *Bgb Bro* double mutants. After several unsuccessful attempts, I creating a null *Bgb* allele using Crispr technology, and attempted to analyze germ line clones containing the Bgb^{8} ΔBro^{1} mutations but failed to recover any eggs. This result was unexpected as the strong loss-of-function $Bgb^{KG03779}$ alleles produced viable germline clones and because Bro is not normally expressed in the female germline. Unfortunately because of the sterility of the Bgb^{8} ΔBro^{1} clones I cannot yet genetically address the question of whether there is a requirement for CBF-β function at *SxlPe*.

While my finding that Bgb^{8} ΔBro^{1} mutant germlines produce no eggs was discouraging, there may be a relatively simple solution that will allow my lab to finally address the question of CBF-β function. There are two likely reasons for the sterility of

Bgb^{8} Δ*Bro*^{1} germline clones. The first is that *Bgb* and *Bro* may function redundantly in the female germline and *Bro* can supply the CBF-β function needed for oogenesis if *Bgb* is absent. The second possibility is that *Bgb* alone may be required in the germline and that our results with *Bgb*^{KG03779} were misleading because the *Bgb*^{KG03779} mutant retains sufficient *Bgb* function to produce eggs. The solution to the first problem is to try inducing the *Bgb*^{8} Δ*Bro*^{1} germline clones in the presence of an unlinked *Bro*$^{+}$ transgene to rescue the female germline defect. Since *Bro* mRNA is not detectable in eggs, such rescued clones should allow us to address CBF-β function in the early embryo. If *Bgb*^{KG03779} is partially functional in the germline, we could induce a new *Bro* mutation on the *Bgb*^{KG03779} chromosome using Crispr and analyze clones of the new double mutant. This would not represent the true null condition for both CBF-β proteins; however, it would allow us to assess their roles under conditions in which Bro was completely absent and Bgb function significantly compromised.

Methods

Fly stocks and mutants

Flies were grown at 25C on a standard cornmeal and molasses medium. *w1118*, 56499 (y^{1} w^{1} f^{1} run^{3}/FM7a/Dp(1;Y)y^{+}mal^{102}), 2024, 2139, hsFLP and Bgb^{KG03779} were from http://flybase.bio.indiana.edu.

In situ hybridization

Embryos were collected 3hr 30 minutes after the egg laying, fixed in 10% formaldehyde, 1XPBS, 50mM EGTA and heptane for 50 minutes. *in situ* was performed

on whole mount embryos following previously described protocol (Erickson and Cline, 1993). Template for in vitro RNA transcription was made by PCR amplification with a forward primer and a reverse primer along with T3 promoter using genomic DNA from *w1118* flies. A Digoxygenin labeled antisense RNA probe was synthesized using *in vitro* transcription kit (MAXISCRIPT T3 kit, Ambion) Probe was detected using anti-Digoxygenis antibody (Roche) that cross react with NBT-BCIP solution staining the embryos.

Embryos are mounted in 70% glycerol/1X PBS. Stages of embryo were detected based on number of nuclei that were stained with DAPI by fluorescence. (Foe and Albert 1983). Early and late stages of cycle 13 were detected by the shape of the peripheral nuclei being horizontal in early cycle compared to circular towards late cycle. In cycle 14 every 5-10 minutes is differentiated based on the cellular furrows and nuclear shape being elongated with time.

Following primers are used for PCR amplification,

Sxl probe is 414 bp. 1:2000 dilution,Hybridization at 55ºC.

Forward primer : CCCTACGTCGACGGCATTGCAGC

Reverse primer : TAATACGACTCACTATAGG-GAATGACCCAATGGAATCG

Runt probe 185 bp. 1:500 dilution. Hybridization at 65ºC.

Forward primer : AACGACGAAAACTACTGCGGCG

Reverse primer : AATTAACCCTCACTAAAACGGTCACCTTGATGGCTTTGC

ftz probe 200bp

Forward primer : AATCAGGAGCAGGTGACCAC

Reverse primer : AATTAACCCTCACTAAAGCGGTTGTGTAGAAATAGTCGG

Bgb probe : 237 bp. 1:100 dilution

Forward primer : TATCGGGAACGAAGCATC

Reverse primer : AATTAACCCTCACTAAAGGTCTATCCATCCACGGAAC

Bro probe 246 bp. 1:100 dilution

Forward primer : GCATCATCACCAGAATCTCG

Reverse primer : AATTAACCCTCACTAAAGTCGTTGACAAACCGCATC

Runt transgenic lines

runt·WRPY, runt wild type transgene is 10050bp with -4,702 kb upstream and +2,137 kb downstream sequence of *runt* transcription unit. It also contains two genes *CR43862* and *CR45610* in upstream whose function is unknown. Fragments were amplified using Expand Long Template PCR System (Roche) and cloned into pCR II-TOPO TA vector (Invitrogen). r*unt·WRPY* construct in TA vector was modified to construct *runt·*Δ*WRPY* by deleting the four amino acids WRPY and *runt·WRPW* was generated by changing Y into W using site specific mutagenesis kit (which company).

All the 'Runt' domain mutations Arg-80-Alanine, Arg-139-Alanine, Arg-142-Alanine, Arg-174-Alanine, Arg-177-Alanine mutants were generated in a similar fashion. Sequencing is performed to confirm the WT and the modifications.

All different *runt* constructs were sub cloned from TA vector to pattB transformation vector (7418bp) that was kindly provided by Johannes Bischof, Basler lab, Zurich. Transgenic injections was carried out by Genetic services Inc. MA. Construct were inserted into fly genomic attP2 site on the third chromosome by targeted φC31 mediated specific insertion. All the transgenes were brought into $runt^{3}$/Balancer background and maintained with two copies.

Primers used to generate *runt·WRPY* fragment

Forward : GGAAAAGTGTGTGGAAAACGGTGGA

Reverse: ccaattCCTAGGGGCATCTGATCCCCAAAAATCTGG

Mutagenesis primers used to *runt·*Δ*WRPY* fragment

Forward : GCACACAGGCGCGCCTTTAAG

Reverse : ccaattCCTAGGCTACACGGTCTTCTGCTGCACGGC

Mutagenesis primers used to generate *runt·WRPW* fragment

Forward : GCACACAGGCGCGCCTTTAAG

Reverse: ccaattCCTAGGCTACCAGGGCCGCCACACGGTC

Mutagenesis primers used to generate ‘Runt’ domain modifications.

Runt-C127S

Forward: CCTCGATCCTCAGCAGCGCCCTG

Reverse:GCGATCCGGTCTGGGCCAGCTC

Runt-K199A

Forward:CTGACCATCACCATTGCCACCTATCCG

Reverse:CGTGAAGGAAGCGCCGCGTCC

Runt-R80A-F

Forward : CCCAATCACTGGGCGTCGAACAAG

Reverse: CAGGGCGCTGCAGAGGATCGAG

Runt-R174A

Forward: TCGACGGGCCAGCGGAGC

Reverse:CGGTCACCTTGATGGCTTTGCTG

Runt-R177A

Forward: GGAGCCAGCAAGTAAGCAAAGTG

Reverse: CGTGGCCCGTCGACG

Runt-R139A

Forward: GGCAAGTCCTTCACGCTGACCATC

Reverse: GCGTCCCGATGCTCCCACG

Runt-R142A

Forward: CCTTCACGCTGACCATCACCATTGC

Revers: ACTTGCCGGCTCCCGATCGTC

1.4kb *SxlPe-LacZ* transgenic line

A 1.4kb *SxlPe* fragment was made by PCR amplification from flies genomic DNA and clone into pCR II-TOPO TA vector (Invitrogen later subcloned to pattB-*LacZ* reporter vector. It was inserted into fly genome at attP2 region by targeted φC31 mediated transgenesis.

A single Runt binding sequence 'TGCGGC' is inserted site at -546, '2-site transgene' with two sites at -546 and -750, '3-site transgenes' with three binding sites -546, -750 and -755, 'tandem repeat transgene' with 3(TGCGGC) at-540 were generated by site specific sequential mutagenesis.

Generation of *Bro* null allele by targeted ends-out site directed mutagenesis strategy

Bro knock-out was generated by deleting a +546 bp coding region along with -86 bp of the 5' UTR to ensure a complete loss of function and this region is replaced by a

w^+ gene that serves as a genetic marker. The 'knock-out' construct was generated using *P*-element vector pW25 (kindly provided by Magger KA) containing 4 kb sequence flanking on either side of the *Bro* region to be deleted with FRT sequences cloned in the same direction. A 'donor' transgenic line was generated by introducing pW25 into fly genome. Its insertion was confirmed by PCR and the presence of a w^+ marker that gives a light red eye in a white eye background. The donor transgene was brought into the flies with hs-FLP by genetic crosses. When larvae express FLPase enzyme and I-SceI enzymes after the heat shock a circular fragment from the donor is excised and linearized. This linear fragment replaces the endogenous *Bro* gene with the deleted version by homologous recombination at the flanking region and also introduce w^+ marker in the germ cells. Progeny of these flies are screened for red eye and confirmed that the eye color retain after induction of FLPase eliminating the possibility of false positives. An extensive PCR was performed to confirm the location of w^+ marker replacing the *Bro* leading to a deletion thus generating a null allele (ΔBro^1).

Germline clones by dominant female sterile technique (DFS)

Bgb loss of function allele $Bgb^{KG03779}$ (Bgb^p)was recombined with FRT onto the proximal region to generate FRT- Bgb^p chromosome. It is brought into hsFLP fly stock that is driven by heat shock promoter generating the stock, hsFLP/hsFLP; FRT-Bgb^p/TM3(Sb). Females from this stock were crossed with males carrying

Ovo^{D1} /TM3(Sb) that is a female dominant sterile allele containing FRT site at the same location as it is on the FRT-*Bgb*^{p}chromosome. Progeny of this cross will have the hsFLP/+; FRT-*Bgb*^{p}/FRT- Ovo^{D1} and females with this genotype are sterile. When an FLPase enzyme is available as larvae are heat shocked during early developmental stages at the time when germline is forming induces recombination between the homologous chromosomes at FRT. In response germ cells generate *Bgb*^{p} homozygous cells in germ line. Ovo^{D1} being dominant female sterile, only the germ cells homozygous for *Bgb*^{p} produce eggs that devoid of Bgb protein.Fig.4.2.

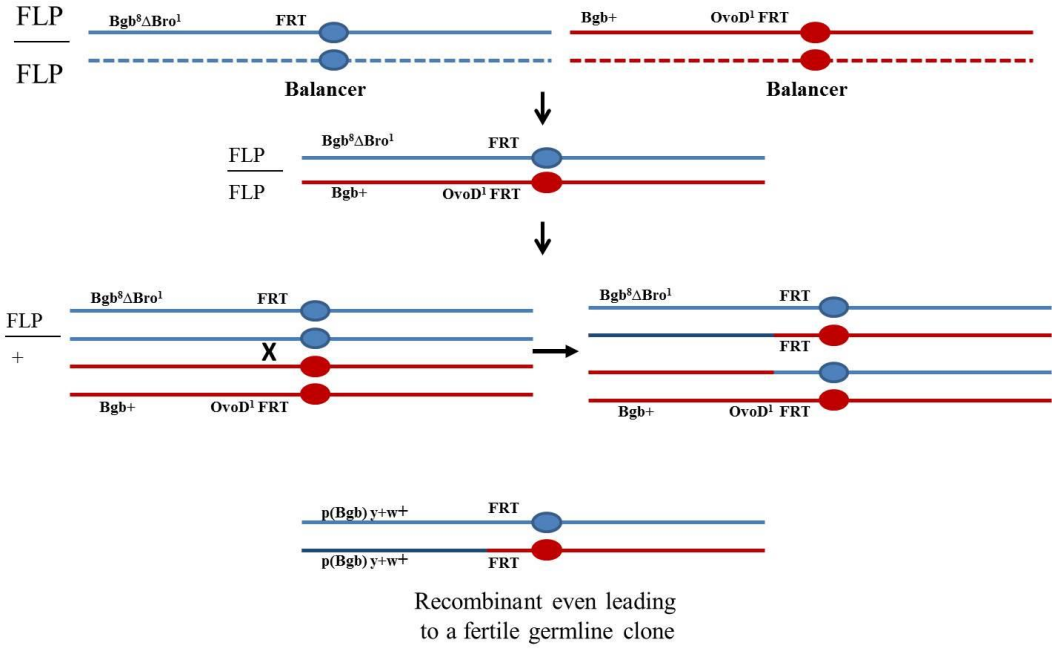

Fig.4.2. Generating germline clones of *Bgb*^{8} Δ*Bro*^{1} using dominant female-sterile technique.

Female larvae carrying the Bgb^{8} ΔBro^{1} / Ovo^{D1} chromosomes along with the FLPase were heat shocked to induce FLPase expression. Specific recombination catalyzed by FLPase leads to homozygous Bgb^{8} ΔBro^{1} . Germ cells with only the double mutants are fertile while the dominant Ovo^{D1} mutant containing germ cells are sterile.

Male recombination

Recombination is induced in males at the p-element of Bgb^{p} allele by providing transposase to recombine with ΔBro^{1} null allele chromosome Bgb^{p} is marked with y+ w+ and the chromosome also has a *Pr* dominant marker on the 3R. Production transposase induce recombination and male progeny with y+ w+ devoid of *Pr* were considered candidates. 49 such candidates were recovered and were checked for the presence of Bgb^{p} by crossing back with Bgb^{p} stocks as they were expected to show homozygous lethality.

EMS mutagenesis screen

100 ΔBro^{1}males were starved overnight for 12 hrs. 25nM solution was made in 1% sucrose solution. 2ml was dispensed into the bottle with filter paper and flies. Flies were left in the bottles for 24 hrs. and transferred into a fresh bottle with food. They were let to recover for 24 hrs. later individual males were crossed with several virgin females. In order to avoid the mosaicism we generated F2 progeny to screen for the mutations that avoid false positives. Finally 5,000 single males were crossed in order to

find the possible mutant. Out of 6 candidates that showed lethality with Bgb^{p} allele showed no sequence change in *the Bgb* region.

CRISPR targeting to generate *Bgb Bro* double mutant

We injected the plasmids with a short sgRNA with *Bgb* guide sequence that are under the control of *U6b* promoter to the flies that constantly express Cas9 in their germ line driven by *nos* promoter in ΔBro^{1} background.

Bgb target sequence: AGGTTCGGTACACGGGCTAT

Both *Bgb* and *Bro* were targeted in the flies that constantly express Cas9 in their germ line driven by *vas* promoter in ΔBro^{1} background. Injection service was provided by Best gene. Inc. *Bgb* target was same as mentioned above.

Bro target sequence: GGTTCGCTACACAGGATACCGGG

REFERENCES

1. Lewis, J, From signals to patterns: space, time, and mathematics in developmental biology. Science, 2008. 322(5900): p. 399-403.

2. Flores-saaib, R. and A. Courey, Regulation of dorso/ventral patterning in the *Drosophila* embryo by multiple dorsal-interacting proteins. Cell Biochemistry and Biophysics, 2000. 33(1): p. 1-17.

3. Grimm, O., M. Coppey, and E. Wieschaus, Modelling the Bicoid gradient. Development, 2010. 137(14): p. 2253-64.

4. Chopra, V.S. and M. Levine, Combinatorial patterning mechanisms in the *Drosophila* embryo. Brief Funct Genomic Proteomic, 2009. 8(4): p. 243-9.

5. Meyer, B.J, Targeting X chromosomes for repression. Current Opinion in Genetics & Development, 2010. 20(2): p. 179-189.

6. Bridges, C.B, Non-disjunction as proof of the chromosome theory of heredity. Genetics, 1916. 1: p. 1-52.

7. Erickson, J.W. and J.J. Quintero, Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. PLoS Biol, 2007. 5(12): p. e332.

8. Cline, T.., Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. Genetics, 1984. 107: p. 231-277.

9. González, A.N., H. Lu, and J.W. Erickson, A shared enhancer controls a temporal switch between promoters during *Drosophila* primary sex determination. Proceedings of the National Academy of Sciences, 2008. 105(47): p. 18436-18441.

10. Keyes, L.N., T.W. Cline, and P. Schedl, The primary sex determination signal of *Drosophila* acts at the level of transcription. Cell, 1992. 68: p. 993-943.

11. Cline, T.W. and B.J. Meyer, Vive la difference: males vs females in flies vs worms. Annu Rev Genet, 1996. 30: p. 637-702.

12. Lucchesi, J.C., W.G. Kelly, and B. Panning, Chromatin remodeling in dosage compensation. Ann Rev Genet, 2005. 39: p. 615-651.

13. Sanchez, L., B. Granadino, and M. Torres, Sex determination in *Drosophila melanogaster*: X-linked genes involved in the initial step of *Sex-lethal* activation. Dev. Genet., 1994. 15: p. 251-264.

14. Penalva, L.O. and L. Sanchez, RNA binding protein sex-lethal (Sxl) and control of *Drosophila* sex determination and dosage compensation. Microbiol Mol Biol Rev, 2003. 67(3): p. 343-59, table of contents.

15. Bell, L.R., J. I. Horabin, P. Schedl and T. W. Cline, Positive autoregulation of *Sex-lethal* by alternative splicing maintains the female determined state in *Drosophila*. Cell, 1991. 65: p. 229-239.

16. Nagengast, A.A., S. M. Stitzinger, C. H. Tseng, S. M. Mount. and H. K. Salz, Sex-lethal splicing autoregulation in vivo: interactions between Sex-lethal, the U1 snRNP and U2AF underlie male exon skipping. Development, 2003. 130(3): p. 463-71.

17. Valcarcel, J., R. Singh, P.D. Zamore, and M.R. Green, The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature, 1993. 362(6416): p. 171-175.

18. Baker, B.S., Sex in flies: the splice of life. Nature, 1989. 340(6234): p. 521-524.

19. Heinrichs, V., L.C. Ryner, and B.S. Baker, Regulation of sex-specific selection of fruitless 5′ splice sites by transformer and transformer-2. Molecular and Cellular Biology, 1998. 18(1): p. 450-458.

20. Demir, E. and B.J. Dickson, Fruitless splicing specifies male courtship behavior in *Drosophila*. Cell, 2005. 121(5): p. 785-94.

21. Gebauer, F., M. Grskovic, and M.W. Hentze, *Drosophila* Sex-Lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Molecular Cell, 2003. 11(5): p. 1397-1404.

22. Duncan, K.E., C. Strein, and M.W. Hentze, The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA. Mol Cell, 2009. 36(4): p. 571-82.

23. Duffy, J.B. and J.P. Gergen, The *Drosophila* segmentation gene *runt* acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene *Sex-lethal*. Genes Dev., 1991. 5: p. 2176-2187.

24. Erickson, J.W. and T.W. Cline, Molecular nature of the *Drosophila* sex determination signal and its link to neurogenesis. Science, 1991. 251: p. 1071-1074.

25. Jinks, T.M., A.D. Polydorides, G. Calhoun. and P. Schedl, The JAK/STAT signaling pathway is required for the initial choice of sexual identity in *Drosophila melanogaster*. Mol Cell, 2000. 5(3): p. 581-7.

26. Sefton, L., J. R. Timmer, Y. Zhang, F. Beranger and Cline, T. W, An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element. Nature, 2000. 405(6789): p. 970-3.

27. Walker, J.J., K.K. Lee, R.N. Desai and Erickson, J. W, The *Drosophila melanogaster* sex determination gene *sisA* is required in yolk nuclei for midgut formation. Genetics, 2000. 155(1): p. 191-202.

28. Kappes, G., G. Deshpande, B.B Mulvey, J.I. Horabin and Schedl, P, The *Drosophila Myc* gene, *diminutive*, is a positive regulator of the *Sex-lethal* establishment promoter, *Sxl-Pe*. Proceedings of the National Academy of Sciences, 2011. 108(4): p. 1543-8.

29. Lott, S.E., Villalta, J. E., Schroth, G. P., Luo, S, Tonkin, L.A., et al., Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-Seq. PLoS Biol, 2011. 9(2): p. e1000590.

30. Yang, D., H. Lu, Y. Hong, T.M. Jinks, P.A. Estes, et al., Interpretation of X chromosome dose at *Sex-lethal* requires non-E-box sites for the basic helix-loop-helix proteins SISB and daughterless. Mol Cell Biol, 2001. 21(5): p. 1581-92.

31. Cline, T.W., Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X:A sex determination signal in *Drosophila* that switch *Sex-lethal* between two alternative stable expression states. Genetics, 1988. 119: p. 829-862.

32. Erickson, J.W. and T.W. Cline, A bZIP protein, SISTERLESS-A, collaborates with bHLH transcription factors early in *Drosophila* development to determine sex. Genes Dev., 1993. 7(9): p. 1688-1702.

33. Avila, F.W. and J.W. Erickson, *Drosophila* JAK/STAT pathway reveals distinct initiation and reinforcement steps in early transcription of *Sxl*. Current Biology, 2007. 17(7): p. 643-648.

34. Klingler, M., J. Soong, B. Butler and J.P. Gergen, Disperse versus compact elements for the regulation of *runt* stripes in *Drosophila*. Dev Biol, 1996. 177(1): p. 73-84.

35. Kramer, S.G., T.M. Jinks, P. Schedl and J.P. Gergen, Direct activation of *Sex-lethal* transcription by the *Drosophila* runt protein. Development, 1999. 126(1): p. 191-200.

36. Adya, N., L.H. Castilla, and P.P. Liu, Function of CBFβ/Bro proteins. Semin Cell Dev Biol, 2000. 11(5): p. 361-368.

37. Fujioka, M., G.L. Yusibova, C.M. Sackerson, S. Tillib, A. Mazo, et at., Runt domain partner proteins enhance DNA binding and transcriptional repression in cultured *Drosophila* cells. Genes Cells, 1996. 1(8): p. 741-54.

38. Golling, G., L. Li, M. Pepling, M. Stebbins, and J.P. Gergen, *Drosophila* homologs of the proto-oncogene product PEBP2/CBF beta regulate the DNA-binding properties of Runt. Mol Cell Biol, 1996. 16(3): p. 932-42.

39. Tsai, C. and J.P. Gergen, Gap gene properties of the pair-rule gene *runt* during *Drosophila* segmentation. Development, 1994. 120(6): p. 1671-83.

40. Bier, E., H. Vaessin, S. Younger-Shepherd, S.L. Jan, and Y.N. Jan, *deadpan*, an essential pan-neural gene in *Drosophila*, encodes a helix-loop-helix protein similar to the hairy gene product. Genes Dev., 1992. 6: p. 2137-2151.

41. Younger-Shepherd, S., H. Vaessin, E. Bier, L. Y. Jan, and Y.N. Jan, *deadpan* an essential pan-neural gene encoding an HLH protein, acts as a denominator element in *Drosophila* sex determination. Cell, 1992. 70: p. 911-921.

42. Barbash, D.A. and T.W. Cline, Genetic and molecular analysis of the autosomal component of the primary sex determination signal of *Drosophila melanogaster*. Genetics, 1995. 141(4): p. 1451-71.

43. Estes, P.A., L.N. Keyes, and P. Schedl, Multiple response elements in the *Sex-lethal* early promoter ensure its female-specific expression pattern. Mol. Cell Biol., 1995. 15: p. 904-917.

44. Lu, H., E. Kozhina, S. Mahadevaraju, D. Yang, F.W. Avila, et al., Maternal Groucho and bHLH repressors amplify the dose-sensitive X chromosome signal in *Drosophila* sex determination. Developmental Biology, 2008. 323(2): p. 248-260.

45. Chen, G. and A.J. Courey, Groucho/TLE family proteins and transcriptional repression. Gene, 2000. 249(1-2): p. 1-16.

46. Jennings, B.H., S.M. Wainwright, and D. Ish-Horowicz, Differential in vivo requirements for oligomerization during Groucho-mediated repression. EMBO Rep, 2008. 9(1): p. 76-83.

47. Li, L.M. and D.N. Arnosti, Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. Current Biology, 2011. 21(5): p. 406-412.

48. Cinnamon, E., A. Helman, S.R. Ben-Haroush, A. Orian, G. Jimenez, et al., Multiple RTK pathways downregulate Groucho-mediated repression in *Drosophila* embryogenesis. Development, 2008. 135(5): p. 829-37.

49. Otto, F., M. Lubbert, and M. Stock, Upstream and downstream targets of RUNX proteins. J Cell Biochem, 2003. 89(1): p. 9-18.

50. Ito, Y., Oncogenic potential of the *RUNX* gene family: /`Overview/'. Oncogene, 2000. 23(24): p. 4198-4208.

51. Cameron, E.R., K. Blyth, L. Hanlon, A. Kilbey, N. Mackay, et al., The *Runx* genes as dominant oncogenes. Blood Cells, Molecules, and Diseases, 2003. 30(2): p. 194-200.

52. Blyth, K., E.R. Cameron, and J.C. Neil, The RUNX genes: gain or loss of function in cancer. Nat Rev Cancer, 2005. 5(5): p. 376-87.

53. Chuang, L.S., K. Ito, and Y. Ito, RUNX family: Regulation and diversification of roles through interacting proteins. Int J Cancer, 2013. 132(6): p. 1260-71.

54. Nusslein-Volhard, C. and E. Wieschaus, Mutations affecting segment number and polarity in *Drosophila*. Nature, 1980. 287(5785): p. 795-801.

55. Ogawa, E., M. Maruyama, H. Kagoshima, M. Inuzuka, J. Lu, et al.,PEBP2/PEA2 represents a family of transcription factors homologous to the products of the *Drosophila runt* gene and the human AML1 gene. Proceedings of the National Academy of Sciences, 1993. 90(14): p. 6859-6863.

56. Aronson, B.D., A.L. Fisher, K. Blechman, M.Caudy and J.P. Gergen, Groucho-dependent and -independent repression activities of Runt domain proteins. Mol Cell Biol, 1997. 17(9): p. 5581-7.

57. Swantek, D. and J.P. Gergen, *Ftz* modulates *Runt*-dependent activation and repression of segment-polarity gene transcription. Development, 2004. 131(10): p. 2281-90.

58. Wheeler, J.C., C. VanderZwan, X. Xu, D. Swantek, W.D. Tracey, et al., Distinct in vivo requirements for establishment versus maintenance of transcriptional repression. Nat Genet, 2002. 32(1): p. 206-10.

59. Li, L.H. and J.P. Gergen, Differential interactions between Brother proteins and Runt domain proteins in the *Drosophila* embryo and eye. Development, 1999. 126(15): p. 3313-22.

60. Kaminker, J.S., R. Singh, T. Lebestky, H. Yan and U. Banerjee, Redundant function of Runt domain binding partners, Big brother and Brother, during *Drosophila* development. Development, 2001. 128(14): p. 2639-2648.

61. Yokomizo, T., M. Yanagida, G. Huang, M. Osato, C. Honda, et al., Genetic evidence of PEBP2beta-independent activation of Runx1 in the murine embryo. Int J Hematol, 2008. 88(2): p. 134-8.

62. Torres, M. and L. Sanchez, The segmentation gene *runt* is needed to activate *Sex-lethal*, a gene that controls sex determination and dosage compensation in *Drosophila*. Genet. Res.,Camb,, 1992. 59: p. 189-198.

63. Paroush, Z., R.L.Jr. Finley, T. Kidd, S.M. Wainwright, P.W. Ingham, et al., Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell, 1994. 79: p. 805-815.

64. Riz, I., H.J. Lee, K.K. Baxter, R. Behnam, T.S Hawley, et al., Transcriptional activation by TLX1/HOX11 involves Gro/TLE corepressors. Biochemical and Biophysical Research Communications, 2009. 380(2): p. 361-365.

65. Butler, B.A., J. Soong, and J. P. Gergen, The *Drosophila* segmentation gene *runt* has an extended cis-regulatory region that is required for vital expression at other stages of development. Mechanisms of Development, 1992. 39(1–2): p. 17-28.

66. Hafen, E., A. Kuroiwa, and W.J. Gehring, Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. Cell, 1984. 37(3): p. 833-841.

67. Jiménez, G., S.M. Pinchin, and D. Ish-Horowicz, In vivo interactions of the *Drosophila Hairy* and *Runt* transcriptional repressors with target promoters. The EMBO Journal, 1996. 15(24): p. 7088-7098.

68. Bravo, J., Z. Li., N. A. Speck. And A. J. Warren, The leukemia-associated AML1 (Runx1)–CBFβ complex functions as a DNA-induced molecular clamp. Nature Structural Biology, 2001. 8(4): p. 371-378.

69. Ogawa, E., M. Inuzuka, M. Maruyama, M. Satake, M. Naito-Fujimoto, et al., Molecular cloning and characterization of PEBP2β, the heterodimeric partner of a novel *Drosophila* runt-related DNA binding protein PEBP2α. Virology, 1993. 194(1): p. 314-331.

70. Kagoshima, H., Y. Akamatsu., Y. Ito., K. Shigesada., et al., Functional dissection of the α and β subunits of transcription factor PEBP2 and the redox susceptibility of its DNA binding activity. The American Society for Biochemistry and Molecular Biology, 1996. 271(51): p. 33074-33082.

71. Tang,Y., B.E. Crute, I. Kelley, J. John, X. Huang, J. Yan, et al., Biophysical characterization of interactions between the core binding factor α and β subunits and DNA. FEBS Letters, 2000. 470(2): p. 167-172.

72. Nagata, T., V. Gupta, D. Sorce, W. Kim, A. Sali, et al., Immunoglobulin motif DNA recognition and heterodimerization of the PEBP2/CBF Runt domain. Nat Struct Mol Biol, 1999. 6(7): p. 615-619.

73. Ito, Y., S.-C. Bae, and L.S.H. Chuang, The RUNX family: developmental regulators in cancer. Nat Rev Cancer, 2015. 15(2): p. 81-95.

74. Gong, W.J. and K.G. Golic, Ends-out, or replacement, gene targeting in *Drosophila*. Proceedings of the National Academy of Sciences, 2003. 100(5): p. 2556-2561.

75. Maggert, K.A., W.J. Gong, and K.G. Golic, Methods for homologous recombination in *Drosophila*. Methods Mol Biiol, 2008. 420(155-174).

76. Preston, C.R. and W.R. Engels, P-Element-induced male recombination and gene conversion in *Drosophila*. Genetics, 1996. 144.

77. Bassett, A.R., C. Tibbit, C.P. Ponting and J.L. Liu, Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Rep, 2013. 4(1): p. 220-8.

78. Bassett, A.R. and J.-L. Liu, CRISPR/Cas9 and genome editing in *Drosophila*. Journal of Genetics and Genomics, 2014. 41(1): p. 7-19.

79. Ren, X., J. Sun, B.E. Housden, Y. Hu, C. Roesel, et al., Optimized gene editing technology for *Drosophila melanogaster* using germ line-specific Cas9. Proceedings of the National Academy of Sciences, 2013. 110(47): p. 19012-19017.

80. Möller, A., F.W. Avila, J.W. Erickson and H. Jäckle, *Drosophila* BAP60 is an essential component of the Brahma complex, required for gene activation and repression. Journal of Molecular Biology, 2005. 352(2): p. 329-337.

81. Ni, Z.,E. Karaskov, T. Yu, S.M. Callaghan, S. Der, et al., Apical role for BRG1 in cytokine-induced promoter assembly. Proceedings of the National Academy of Sciences, 2005. 102(41): p. 14611-14616.

82. Ni, Z. and R. Bremner, Brahma-related gene 1-dependent STAT3 recruitment at IL-6-inducible genes. The Journal of Immunology, 2006. 178(1): p. 345-351.

83. Zhang, Y., M. Cheng, Y. Zhang, X. Zhong, H. Dai, et al., A switch from hBrm to Brg1 at IFN[gamma]-activated sequences mediates the activation of human genes. Cell Res, 2010. 20(12): p. 1345-1360.

84. Salz, H.K. and J.W. Erickson, Sex determination in *Drosophila*: the view from the top. Fly, 2010. 4(1): p. 60-70.

85. Kaul, A.K., E.F. Schuster, and B.H. Jennings, Recent insights into *Groucho* co-repressor recruitment and function. Transcription, 2015: p. e1000709.

86. Lee, W., B.C. Andrews, M. Faust, U. Walldorf, and E.M. Verheyen, Hipk is an essential protein that promotes Notch signal transduction in the *Drosophila* eye by inhibition of the global co-repressor *Groucho*. Dev Biol, 2009. 325(1): p. 263-72.

87. Wee, H.-J., D.C. Voon, S. Bae and Y. Ito,PEBP2-β/CBF-β–dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. Blood, 2008. 112(9): p. 3777-3787.

88. Blythe, Shelby A. and Eric F. Wieschaus, Zygotic genome activation triggers the DNA replication checkpoint at the midblastula transition. Cell, 2015. 160(6): p. 1169-1181.